BARD Final Report IS-4801-15C

Uncovering rumen microbiome components shaping feed efficiency in dairy cows

Mizrahi, I. Agricultural Research Organization
White, B.A. Board of Trustees, Univ. of Illinois

Project award year: 2015
Three year research project

BARD Project No. IS-4801-15C, Final Scientific Report
Uncovering rumen microbiome components shaping feed efficiency in dairy
Prof. Itzhak Mizrahi (BGU, Israel); Prof Bryan White (CCBGM, USA)

Ruminants provide human society with high quality food from non-human-edible resources, but their emissions negatively impact the environment via greenhouse gas production. The rumen and its resident microorganisms dictate both processes. The overall goal of this project was to determine whether a causal relationship exists between the rumen microbiome and the host animal's physiology, and if so, to isolate and examine the specific determinants that enable this causality. To this end, we divided the project into three specific parts: (1) determining the feed efficiency of 200 milking cows, (2) determining whether the feed-efficiency phenotype can be transferred by transplantation and (3) isolating and examining microbial consortia that can affect the feed-efficiency phenotype by their transplantation into germ-free ruminants.

We finally included 1000 dairy cow metadata in our study that revealed a global core microbiome present in the rumen whose composition and abundance predicted many of the cows' production phenotypes, including methane emission. Certain members of the core microbiome are heritable and have strong associations to cardinal rumen metabolites and fermentation products that govern the efficiency of milk production. These heritable core microbes therefore present primary targets for rumen manipulation towards sustainable and environmentally friendly agriculture. We then went beyond examining the metagenomic content, and asked whether microbes behave differently with relation to the host efficiency state. We sampled twelve animals with two extreme efficiency phenotypes, high efficiency and low efficiency where the first represents animals that maximize energy utilization from their feed whilst the later represents animals with very low utilization of the energy from their feed. Our analysis revealed differences in two host efficiency states in terms of the microbial expression profiles both with regards to protein identities and quantities.

Another aim of the proposal was the cultivation of undescribed rumen microorganisms is one of the most important tasks in rumen microbiology. Our findings from phylogenetic analysis of cultured OTUs on the lower branches of the phylogenetic tree suggest that multifactorial traits govern cultivability. Interestingly, most of the cultured OTUs belonged to the rare rumen biosphere. These cultured OTUs could not be detected in the rumen microbiome, even when we surveyed it across 38 rumen microbiome samples. These findings add another unique dimension to the complexity of the rumen microbiome and suggest that a large number of different organisms can be cultured in a single cultivation effort.

In the context of the grant, the establishment of ruminant germ-free facility was possible and preliminary experiments were successful, which open up the way for direct applications of the new concepts discovered here, prior to the larger scale implementation at the agricultural level.

Summary Sheet

Publication Summary

PubType	IS only	Joint	US only
Reviewed	0	2	0

Training Summary

Trainee Type	Last Name	First Name	Institution	Country
M.Sc. Student	Crooks	Taylor Crooks	University of Illinois	USA
M.Sc. Student	Zorea	Alvah	Ben gurion University	Israel
Postdoctoral Fellow	Bogumil	David	Ben gurion University	Israel
Postdoctoral Fellow	Kugel	Sasha	Ben gurion University	Israel

Contribution of collaboration

Our proposal contained experimental designs based on cooperative endeavors involving each of the partners' expertises. Our study combined the specific expertise of the US group in rumen microbiology, samples sequencing, metagenomic profiling and computational biology with the Israeli group's metabolomic, molecular biology, anaerobic microbiology capabilities and unique germ-free facility and equipment to test a working hypothesis in which feed efficiency of dairy cows would be transmissible via the rumen microbiome. Thus, the proposal was build on the specific strengths of each group. **The collaboration was synergistic in nature and resulted into two-peer reviewed publications in highly ranked scientific journals.**

Achievements

(1) Determining feed efficiency genomic and proteomic parameters

Using our tightly controlled animals cohort, that included Holstein cows sharing highly similar genomic background that were reared in the same facility with identical diet, we revealed that non-only efficient animals contain less rich set of microbes and microbial genes but also that the proteins associated to high feed efficiency were less abundant in number as proteins associated to inefficient state significantly ($p < 1e-6$).

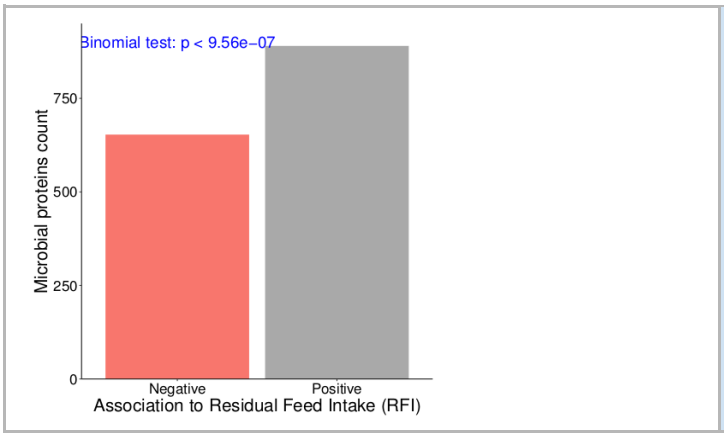

Figure 1: A bar-plot comparing the number of microbial proteins that are typical to efficient animals (negatively correlate with RFI) to those that are typical to inefficient animals (positively correlate to RFI). Color represents efficiency state and Y-axis represents distinct microbial protein count.

Furthermore we revealed that host-state typical proteins were expressed by specific microbial genomes that were taxonomically distant, with Bacteroidetes being linked to efficiency and Firmicutes to inefficiency.

We also identified ecological patterns that clearly separate efficient animals' microbiome from inefficient ones'. We observed that proteins typical to the two states diverge functionally. When we looked into each of the two protein groups, we found that while the efficient microbiome proteins exhibit low functional redundancy, the inefficient microbiomes typical proteins are actually enriched with redundant functions. This implies that niche partitioning on the microbial expression level promotes the overall community efficiency. A logical explanation for this observation could be that lower functional redundancy reflects a more coordinated inter-microbial interaction. With less competition over resources, more

energy could instead be utilized into higher yield of final fermentation products. Such products, such as important SCFA like Propionate and Acetate then affect, in their availability, the growth and yield of the hosting animal.

(2) Determining whether the feed-efficiency phenotype can be transferred by transplantation

Our newly established unique germ-free lamb facility can enable us to directly measure the impact of essential communities on animal physiology. This facility houses eight isolators that can each contain two animals at a time. We validated animal physiology (i.e., weight gain, feed consumption, blood metabolites and immune system profile, using cytokine and immunoglobulin profiling kits), and their rumen will be sampled regularly (using stomach tubes) during the time of animal isolation. In our dietary regime protocols, lambs are fed sterilized milk starter and are exposed to a fiber-based diet (roughage) from the first day after birth. The animals are weaned at age 4 weeks, when they are fed a sterilized plant fiber-based diet. Lamb birth weight is approximately 4 kg, and lambs grown under conditions wherein their natural rumen communities are confined within the isolators gain between 0.125 g to 0.175 g per day, which is similar to natural growth.

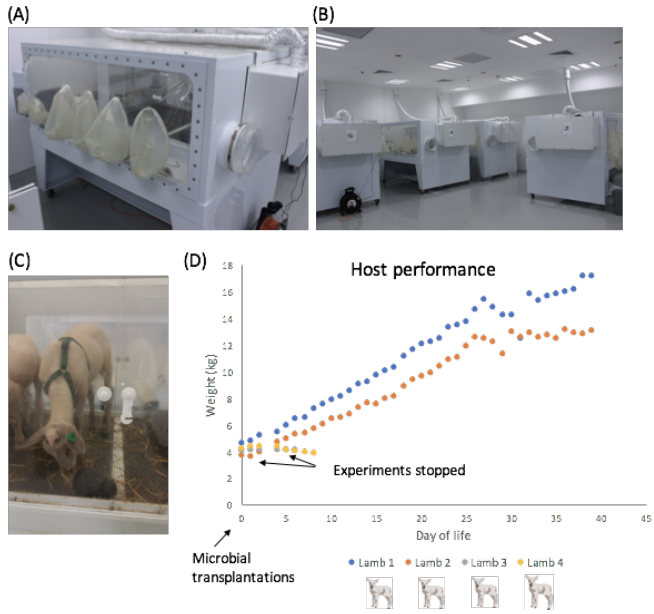

Figure 2: A, B, C, D. Our functioning ruminant germ-free facility. A: View of the exterior of an isolator. B: View of isolators inside the germ-free ruminant facility. C: Lambs with their natural microbiome (non gnotobiotic) inside the isolators. **D. Host health measurements**. At birth, animals were transplanted with microbial communities. Daily weight was monitored to measure life sustainability by the microbiome. Experiments were pursued in Lambs 1 and 2, whose weights increased as expected (0.125 g to 0.175 g a day). Experiments will be stopped in animals whose weight does not increase during three consecutive days (e.g., Lambs 3 and 4), and the animals will be transplanted with full microbiomes.

(3) Isolating and examining microbial consortia that can affect the feed-efficiency phenotype by their transplantation into germ-free ruminants.

During our studies, we asked two fundamental questions: (i) What is the portion of the rumen microbiome that can potentially be cultivated? (ii) How are medium type, sample dilution and phylogeny related to cultivability? To answer these questions, we used defined and undefined (with rumen fluid) anaerobic media and plated decimal dilutions of a rumen sample. Using this methodology, 23% of the rumen microbiome is potentially cultivable. We found a positive correlation between the cultivability of an OTU (operational taxonomic unit) and its abundance in the original rumen sample for both media. The effect of sample dilution exceeded the effect of medium type, indicating that the laborious plating effort is worth the chance of capturing rumen microbial diversity. Furthermore, technical replications in plating strongly contributed to the richness on the plates and both media had a high number of unique OTUs. Cultivation seems to select for a multifactorial set of genetic traits scattered across the phylogenetic tree. Despite the great variety of microbes cultured on the plates, the selection for genomic traits also calls for new isolation approaches based on certain sets of traits to increase cultivated diversity, as these genetic traits are not necessarily interconnected and unlikely to be fully dependent on each other.

The OTUs cultured here were mainly from the rare rumen biosphere. Taken together, the high relevance of the rare biosphere reported in the literature and its surprisingly high cultivability as found here, we concluded that the rumen's complexity and functional repertoire are beyond our current estimation. Nonetheless part of these rare biosphere microbes can be studied in pure culture, enabling a deeper understanding of rumen ecosystem functionality. Therefore, our study presents a very promising notion: a vast variety of unknown rumen microorganisms can be grown using defined and undefined anaerobic media.

Publications for Project IS-4801-15C

Status	Type	Authors	Title	Journal	Vol:pg Year	Coun
Published	Reviewed	*Shabat, Sheerli Kruger Ben Sasson, Goor Doron-Faigenboim, Adi Durman, Thomer Yaacoby, Shamay Berg Miller, Margret E. White, Bryan A. Shterzer, Naama Mizrahi, Itzhak*	Specific microbiome-dependent mechanisms underlie the energy harvest efficiency of ruminants	***ISME J***	10 : 2958-2972 2016	Joint
Published	Reviewed	*Goor Sasson, Sheerli Kruger Ben-Shabat, Eyal Seroussi, Adi Doron-Faigenboim, Naama Shterzer, Shamay Yaacoby, Margret E. Berg Miller, Bryan A. White, Eran Halperin, Itzhak Mizrahi*	Heritable Bovine Rumen Bacteria Are Phylogenetically Related and Correlated with the Cow's Capacity To Harvest Energy from Its Feed	***mBio***	8 : Issue 4 e00703-17 2017	Joint

Appendix

The ISME Journal (2016) 10, 2958–2972
© 2016 International Society for Microbial Ecology All rights reserved 1751-7362/16
www.nature.com/ismej

OPEN

ORIGINAL ARTICLE

Specific microbiome-dependent mechanisms underlie the energy harvest efficiency of ruminants

Sheerli Kruger Ben Shabat[1,2,3], Goor Sasson[1], Adi Doron-Faigenboim[2], Thomer Durman[1], Shamay Yaacoby[2], Margret E Berg Miller[4,5], Bryan A White[4,5], Naama Shterzer[1] and Itzhak Mizrahi[1]

[1]*Department of Life Sciences and the National Institute for Biotechnology in the Negev, Ben-Gurion University of the Negev, Beer-Sheva, Israel;* [2]*Department of Ruminant Science, Institute of Animal Sciences, Agricultural Research Organization, Volcani Center, Bet Dagan, Israel;* [3]*Department of Molecular Microbiology and Biotechnology, The George S. Wise Faculty of Life Science, Tel Aviv University, Ramat-Aviv, Israel;* [4]*Department of Animal Sciences, University of Illinois, Urbana, IL, USA and* [5]*The Institute for Genomic Biology, University of Illinois, Urbana, IL, USA*

Ruminants have the remarkable ability to convert human-indigestible plant biomass into human-digestible food products, due to a complex microbiome residing in the rumen compartment of their upper digestive tract. Here we report the discovery that rumen microbiome components are tightly linked to cows' ability to extract energy from their feed, termed feed efficiency. Feed efficiency was measured in 146 milking cows and analyses of the taxonomic composition, gene content, microbial activity and metabolomic composition was performed on the rumen microbiomes from the 78 most extreme animals. Lower richness of microbiome gene content and taxa was tightly linked to higher feed efficiency. Microbiome genes and species accurately predicted the animals' feed efficiency phenotype. Specific enrichment of microbes and metabolic pathways in each of these microbiome groups resulted in better energy and carbon channeling to the animal, while lowering methane emissions to the atmosphere. This ecological and mechanistic understanding of the rumen microbiome could lead to an increase in available food resources and environmentally friendly livestock agriculture.
The ISME Journal (2016) **10,** 2958–2972; doi:10.1038/ismej.2016.62; published online 6 May 2016

Introduction

Ruminants hold enormous significance for man, as they convert the energy stored in plant biomass polymers, which are indigestible for humans, to digestible food products. Humans domesticated these animals for this purpose in the Neolithic era (Ajmone-Marsan *et al.*, 2010) and have been farming them ever since for the production and consumption of animal protein in the form of meat and milk. In today's extensive production regimes, ruminants consume 30% of the crops grown on earth and occupy another 30% of the earth's land mass (Thornton, 2010). These animals also emit methane —a highly potent greenhouse gas—to the atmosphere and are considered to be responsible for a considerable portion of its emission because of anthropogenic activities (McMichael *et al.*, 2007). One way to tackle these problems is to increase the animals' energetic efficiency, that is, the efficiency with which they convert energy from feed, thereby increasing food availability, while lowering the environmental burden, as these animals would produce more and eat less (Bradford, 1999; Thornton, 2010).

Different methods are used to evaluate an animal's energetic efficiency; of these, the residual feed intake (RFI) method (Koch *et al.*, 1963) is highly accepted and widely used (Herd and Arthur, 2009) as it takes into account growth and body size and is thus suitable for comparisons between animals. This parameter is an estimation of the difference between an animal's actual feed intake and its predicted feed intake based on its production level and body weight. The energetic efficiency varies considerably between different individuals from the same breed. Specific genomic regions, such as one that is suggested to be associated with a role in controlling energy metabolism, have been found to correlate to feed efficiency using genome-wide association studies (Pryce *et al.*, 2012). Nevertheless, only a moderate genetic component (heritability ranging from 0.26 to 0.58) affects energy utilization, as has also been demonstrated by elevation of feed efficiency via selection of animals according to their RFI (Archer *et al.*, 1999; Moore *et al.*, 2009).

Correspondence: I Mizrahi, Department of Life Sciences and the National Institute for Biotechnology in the Negev, Ben-Gurion University of the Negev, PO Box 653, Beer-Sheva 84105, Israel.
E-mail: imizrahi@bgu.ac.il
Received 6 December 2015; revised 2 March 2016; accepted 22 March 2016; published online 6 May 2016

One important factor that could greatly contribute to variations in these animals' feed efficiency is the rumen microbiome. The ability of these animals to digest plant biomass polymers is attributed to this complex microbiome that resides in their upper digestive tract in a compartment termed the rumen (Mizrahi, 2013). The anaerobic environment in the rumen and the highly complex food webs sustained by the rumen microbiome enable the fermentation of plant material into metabolic end products such as short-chain fatty acids (SCFAs) and methane. Although SCFAs are absorbed through the rumen wall and serve to fulfill the animal's energy needs, methane is not absorbed; it is emitted to the atmosphere together with its retained energy, thereby contributing to energy loss from the feed, as well as global warming (Mizrahi, 2011). Microbial composition of the rumen has been described using various techniques (Brulc *et al.*, 2009; Hess *et al.*, 2011; Jami and Mizrahi, 2012; Henderson *et al.*, 2015). In addition, differences between high and low RFI animals have been reported in terms of methane production, as well as of some differences in microbial composition (Nkrumah *et al.*, 2006; Mizrahi, 2011; Hernandez-Sanabria *et al.*, 2012; Jami *et al.*, 2014; Kittelmann *et al.*, 2014; Shi *et al.*, 2014; Wallace *et al.*, 2015). Nevertheless, a comprehensive and thorough understanding of microbiome structure patterns and how to translate them to functionality at the animal level is still lacking.

Here we determined the feed efficiency phenotype of a cohort of 146 animals; this phenotype was mapped to its underlying microbiome determinants by sampling and analyzing rumen fluid and feces from the animals in the two extreme quartiles (78 animals—40 efficient and 38 inefficient; Supplementary Figure S1 and Supplementary Data 1). We characterized the taxonomic composition, genetic functional potential, metabolomic composition and activity of these rumen microbiomes to explore the hypothesis of a link between the animal's feed efficiency and these microbiome components, and to uncover the potential mechanisms that might explain this link. Our data provide detailed and novel insight into the characteristics and components of the rumen microbiome related to feed efficiency—their ecological context, the underlying mechanisms and their potential as markers to predict the feed efficiency phenotype.

Materials and methods

Trial design

The experimental procedures used in this study were approved by the Faculty Animal Policy and Welfare Committee of the Agricultural Research Organization (ARO), approval number IL-386/12, Volcani Research Center, and were in accordance with the guidelines of the National Council for Animal Experimentation.

A total of 146 Holstein Friesian dairy cows were selected for the experiment and housed at the ARO's experimental dairy farm in Bet Dagan, Israel. Cows with history of diseases, miscarriages and twin pregnancies or that were above first trimester were not included in the experiment. The experimental dairy farm is equipped with a facility that is specially designed to individually monitor all of the animal's functions, feed intake and different physiological parameters (Halachmi *et al.*, 1998). The animals were divided into seven groups according to lactation period such that each cow was between 50 and 150 days of lactation when monitored. Each group contained between 19 and 21 cows that were monitored for 42–49 days. The animals were fed *ad libitum* a standard lactating cow diet consisting of 30% roughage and 70% concentrate and had free access to water. The cows were habituated with the aforementioned diet for 3 weeks before the start of the experiment so that they would become accustomed to their individual feeding station.

The following parameters were automatically monitored three times a day during the experiment: dry matter (DM) intake (kg), weight (kg), milk yield (kg), milk lactose, fat and protein (g) and somatic cell count using the Afimilk program (Afimilk Ltd, Kibbutz Afikim, Israel). Milk samples were sent to an authorized milk quality lab (National Service for Udder Health and Milk Quality, Caesarea, Israel) three times for each group to verify the Afimilk program analysis. Body conditioning score was measured once a week by the same person throughout the experiment.

Feed efficiency parameters RFI and conversion ratio were calculated according to National Research Council (2001) formulas. In order to increase the statistical power compared with random sampling, extreme phenotypes sampling approach (Li *et al.*, 2011) was applied. Twelve cows with the most extreme and stable RFI values were selected from each group for rumen fluid sampling, six with low and six with high RFI values. Tukey's test was used to verify that the RFI value of each cow was steady throughout the experiment and significantly different from cows in the reciprocal efficiency group. Overall, 78 cows were chosen for sampling and represented the 25% most efficient and 25% most inefficient animals of the whole cohort ($P<0.0001$; Supplementary Figure S1 and Supplementary Data 1). Our previous work showed that samples from 16 cows are sufficient in order to cover all of the microbial diversity in the bovine rumen on a specific diet (Jami and Mizrahi, 2012). Therefore, this extreme phenotype characterization and sampling approach, together with the large cohort from which the extreme cows were chosen from, ensured increased adequate power to detect microbiome components connected to the host phenotype.

Sample collection

Rumen samples were collected on 3 consecutive days. The cows were sampled 6 h after feeding in which they were not offered feed; 500 ml of rumen contents were collected using a stainless-steel

stomach tube with a rumen vacuum sampler, and pH was immediately determined. Samples for DNA and metabolite extraction were snap frozen in liquid nitrogen and stored at − 80 °C until analysis. Rumen samples for metabolic assays were filtered through six layers of cheese cloth to remove big feed particles, transferred to CO_2-containing bottles and flushed with CO_2 to maintain anaerobic conditions. Immediately after collection, the rumen samples were maintained at 39 °C up to 1 h until use, and processed in the laboratory, located 100 m away.

Fresh fecal samples were obtained three times a day for 4 consecutive days. Samples were immediately frozen at − 20 °C.

In-vitro digestibility assay

The *in-vitro* digestibility of plant cell wall fibers, represented by neutral detergent fiber or total feed polymers (in DM), was determined according to the two-stage technique by Tilley and Terry (1963). Briefly, cows' feed was dried for 72 h in an aerated 60 °C oven and then ground to pass a 1-mm screen. The feed was incubated with rumen fluid and artificial rumen buffer, in sealed glass tubes.

Artificial rumen buffer was formulated as described previously (McDougall, 1948; Tilley and Terry, 1963). Briefly, 100 ml of buffer A (98 g $NaHCO_3$, 93 g $NaH_2PO_4 \cdot 12H_2O$, 5.7 g KCl, 4.7 g NaCl, 1.2 g $MgSO_4 \cdot 7H_2O$, added to double-distilled water (DDW) to final volume of 1 liter) was added to 800 ml DDW. The solution was flushed with CO_2 to reduce the pH to 6.8–7.0. Then, 1 ml of buffer B (40 g $CaCl_2$ added to DDW to final volume of 1 liter), 50 ml of buffer C (30 g NH_4HCO_3 added to DDW to final volume of 1 liter) and 100 µl of buffer D (10 g $MnCl_2 \cdot 4H_2O$, 1 g $CoCl_2 \cdot 6H_2O$, 8 g $FeCl_3 \cdot 6H_2O$ added to DDW to final volume of 1 liter) were added and the buffer was brought to a final volume of 1 liter with DDW.

The tubes were flushed with CO_2 and closed with a unidirectional valve cap, which only allowed emission of gas from the tube. The tubes were incubated for 24 or 48 h at 39 °C and were shaken five times a day, followed by incubation with acid pepsin. At the end of this procedure, the undigested solids were precipitated by centrifugation at 1000 *g* for 10 min and dried in an aerated oven at 60 °C for 72 h. The precipitates were used for residual DM determination by weighing or for residual neutral detergent fiber determination by following the procedure of Van Soest *et al*. (1991). The results are expressed as mean feed digestibility in the rumen from two consecutive sampling days.

In-vivo digestibility

Fecal grab samples were pooled for each cow, dried at 60 °C for 72 h in a forced-air oven and ground to pass a 1-mm screen. The indigestible neutral detergent fiber content was determined in the ration and in the fecal samples according to a previously reported method (Lippke *et al.*, 1986) after incubation with rumen fluid for 72 h and was used as an internal marker for the apparent total-tract DM digestibility analysis. Each cow's *in-vivo* DM and neutral detergent fiber digestibility of the ration was calculated using its average DM intake and fecal output.

In-vitro methane emission assay

Samples were diluted 1:2 (v/v) with artificial rumen buffer. Duplicates of 5 ml aliquots from each diluted sample were transferred to screw-cap glass tubes (ISI, Israel Scientific Instruments Ltd, Petah-Tikva, Israel) suitable for methane measurement using a gas chromatography (GC) system HP-5890 series II, (Hewlett-Packard, Palo Alto, CA, USA) with a FID detector. The samples were incubated at 39 °C for 24 h with 0.5 g DM feed, and then analyzed by GC for methane emission. Samples of 0.5 ml gas from the tube headspace were injected into a 182.88 cm × 0.3175 cm × 2.1 mm packed Supelco analytical-45/60 Molecular sieve 5 A column (Supelco Inc., Bellefonte, PA, USA) with helium carrier gas set to a flow rate of 10 ml min^{-1} and an oven temperature of 200 °C. The oven temperature remained steady for a total run time of 5 min. A standard curve was generated using pure methane gas.

Methane production was quantified for 36 rumen microbiome samples of the most extreme animals of the feed efficiency groups (18 efficient and 18 inefficient), with two biological repeats of each animal.

Identification and quantification of rumen fluid metabolites

Frozen rumen fluid samples were thawed at 25 °C and centrifuged at 10 000 *g* for 15 min. The supernatant was filtered through a sterile 0.45 µm filter (Merck Millipore Ltd., Tullagreen, County Cork, Ireland). Rumen fluid samples were kept on ice during metabolite extraction in the gas chromatography mass spectrometry and GC metabolite identification and quantification pipelines to minimize metabolite degradation.

The rumen samples were analyzed by gas chromatography mass spectrometry for polar metabolites and by GC with a FID detector for SCFAs. The extraction and derivatization protocol for the gas chromatography mass spectrometry analysis was adapted from a previously described method (Saleem *et al.*, 2013). Derivatized extracts were analyzed using an Agilent 5975C GC and an Agilent 7890A MS (Agilent Technologies, Palo Alto, CA, USA) operating in electron impact (EI) ionization mode. Aliquots (1 µl) were injected (splitless) into a 30 m × 0.25 mm × 0.25 µm HP-5MS Ultra Inert column (Agilent Technologies, Berkshire, UK) with helium carrier gas set to a flow rate of 1 ml min^{-1} and initial oven temperature of 70 °C. The oven temperature was held constant at the initial temperature for 2 min, and thereafter increased at 10 °C min^{-1} to a final

temperature of 310 °C, and a final run time of 45 min. Samples were run using full scan in a mass range of 50–500 m/z (1.7 scan s^{-1}) with a detection delay of 4 min. Retention indices were calculated using a C8-C20 alkane standard mixture solution (Sigma-Aldrich, Buchs, Switzerland) as the external standard. Quantification and identification of trimethylsilylated metabolites were performed using the NIST database and high performance liquid chromatography grade standards.

For SCFA identification and quantification, 400 µl of filtered rumen fluid were mixed with 100 µl of 25% metaphosphoric acid solution (w/v in DDW) and vortexed for 1 min. The samples were incubated at 4 °C for 30 min and subsequently centrifuged for 15 min at 10 600 *g*. The supernatant was decanted into new tubes, then 250 µl methyl tert-butyl ether (Sigma-Aldrich) was added and the tubes were vortexed for 30 s. Another cycle of centrifugation was performed for 1 min at 10 600 *g*. The upper phase, which contained methyl tert-butyl ether +SCFAs, was analyzed using an Agilent 7890B GC system (Agilent Technologies, Santa Clara, CA, USA) with a FID detector. The temperatures at the inlet and detector were 250 °C and 300 °C, respectively. Aliquots (1 µl) were injected with a split ratio of 1 : 25 into a 30 m × 0.32 mm × 0.25 µm ZEBRON ZB-FFAP column (Phenomenex, Torrance, CA, USA) with helium carrier gas set to a flow rate of 2.4 ml min^{-1} and initial oven temperature of 100 °C. The oven temperature was held constant at the initial temperature for 5 min, and thereafter increased at 10 °C min^{-1} to a final temperature 125 °C, and a final run time of 12.5 min.

Quantification and identification of metabolites were performed using high performance liquid chromatography-grade standards. All metabolites were normalized to the organic matter content of the rumen fluid they were extracted from. Rumen samples were filtered through a sterile 0.45 µm Supor Membrane filter (PALL Life Sciences Ann Arbor, MI, USA). The organic C in the rumen samples was analyzed with a Formacs, combustion total organic carbon analyzer (Skalar, De Breda, The Netherlands).

Microbial DNA extraction

The rumen microbial fraction was separated according to Stevenson and Weimer (2007), with minor modifications to suit the needs of these experiments as described in Jami *et al*. (2013). The DNA extraction was performed as described by Stevenson and Weimer (2007).

Shotgun DNA sequencing and analysis

Metagenomic DNA libraries were constructed with the TruSeq DNA Sample Prep kit (Illumina, San Diego, CA, USA). Libraries were pooled and sequenced on two lanes for 151 cycles from each end on a HiSeq2500 (Illumina) and processed with Casava 1.8.2 (Illumina). On average, 35 581 041 ± 6 899 269 paired end reads were obtained from each sample and 2 775 321 186 paired end reads were obtained overall. In all, 18.6% of the reads did not pass artifact filtering and trimming using MOCAT pipeline (Kultima *et al*., 2012).

To obtain a more comprehensive metagenome, a joint assembly of all data from the 78 cows was created. This compensated for the lower sequencing depth of each individual sample and any bias caused by assembly of individual samples. Reads from all samples were pooled and assembled into one metagenome using CLC Bio, package CLC Assembly Cell version 3.2.2 (Qiagen, Redwood, CA, USA) with K-mer = 21 and default parameters; 16 784 830 contigs were obtained. A QC pipeline of dereplication and screening for *Bos taurus* reads was performed using the MG-RAST pipeline. No redundancies were found and 0.43% of the contigs were discarded after removing *Bos taurus* contaminants. The phylogenetic origin of each contig was annotated with RefSeq database (Pruitt *et al*., 2007) ($E \leqslant 10^{-5}$) using the MG-RAST pipeline (Meyer *et al*., 2008).

Gene calling was performed on the contigs using FragGeneScan (Rho *et al*., 2010); 21 531 511 genes were identified over all. Each sample's reads were recruited against the overall genes using burrows-wheeler alignment tool (Li and Durbin, 2009) with 98% identity and default parameters; a threshold of one read for gene identification was chosen to include rare genes in the analysis. On average, 52.4% of the reads from each sample were mapped to the obtained genes, without differences between the efficiency groups (Supplementary Figure S2). An average of 4 079 212 genes were identified in each sample. The abundance of a specific gene was calculated by the number of reads uniquely recruited, normalized to the length of the gene and total reads obtained from the sample. The number of genes detected had no dependence on the number of mapped reads (Supplementary Figure S3).

16S ribosomal DNA sequencing and analysis

The 16S V3 region was amplified using the primers 357F 5′-CCTACGGGAGGCAGCAG-3′ and 926R 5′-C CGTCAATTCMTTTRAGT-3′ (Peterson *et al*., 2009). The libraries were pooled and sequenced on one MiSeq flowcell (Illumina) for 251 cycles from each end of the fragments and analyzed with Casava 1.8. Overall, 49 760 478 paired end reads were obtained for all 3 sampling days, with an average of 212 652 reads per sample per day.

Data quality control and analyses were performed using the QIIME pipeline version 1.7.0 (Caporaso *et al*., 2010). Species were defined at 97% identity using UCLUST (Edgar, 2010). Taxonomy assignment of species was performed using BLAST against the 16S rRNA reference database RDP (version 10) (Cole *et al*.,

2003). All singletons and doubletons were removed from the data set, resulting in 85 225 species with an average of 5039 per sample. Species were binned at different taxonomic levels to receive taxon abundances for each phylogenetic level (Supplementary Figure S4 and Supplementary Data 2).

Biodiversity analysis

Within-sample (alpha) diversity was calculated using Shannon index and dominance was determined according to 1–Simpson index (Harper, 1999). The indices of 16S rRNA gene profiles were calculated using bootstrapping with 9999 replicates. Richness of genes and taxa are presented as simple counts of genes and taxa.

Identifying differential species and genes

The statistical significance of differences in species and gene abundance between the efficiency groups was tested by Wilcoxon rank-sum test coupled with a bootstrapping approach adopted from Le Chatelier *et al.* (2013): 70% of the whole sampled cohort was randomly chosen 30 times and significance was determined at $P < 0.05$ with bootstrap = 0.8 as a threshold. This process was repeated with another 30 iterations on the 48 most extreme cows (24 efficient and 24 inefficient). Overall, 18 significantly different species and 34 166 significantly different genes common to all 60 tests were further analyzed. Species and genes that were significantly different were correlated to the RFI parameter using Spearman correlation. Functional annotation of significant genes was achieved using BLASTP with $E \leqslant 10^{-6}$ against KEGG PATHWAY, MODULE, BRITE, GENES and ORTHOLOGY databases (2014; 25% annotation) (Kanehisa *et al.*, 2011). These genes were also blasted against the NR database (Pruitt *et al.*, 2007), and their phylogenetic annotation was determined according to the best hit (BLASTP with $E \leqslant 10^{-6}$; 89% annotation).

Statistical tests and estimation of false discovery rate

Tukey's, Student's *t* and Wilcoxon rank-sum tests were conducted depending on the normality of distribution of the input data. All tests were corrected for false discovery rate using the method described by Benjamini and Hochberg (1995) unless otherwise noted. In permutation *t*-test, significance of the difference between means was inferred by performing *t*-test between the two groups and comparing the resulting *t*-statistic to the *t*-statistics resulting from 9999 permutations of random group assignments (two-tailed, $P < 0.05$) (Davis, 1986). For multiple hypothesis correction, the distribution of *t*-test *P*-values was compared with the lowest *P*-values distribution resulting from 9999 permutations of random group assignments according to Westfall and Young (1993). This procedure was performed using the R bioconductor package multtest (Pollard *et al.*, 2005), function *mt.maxT*, individually for each metabolic or activity test, namely polymers, SCFAs, methane and all other measured metabolites. Variance similarity was tested where required by the statistical test.

Predictions of different physiological parameters

Feature selection of microbial species and genes was conducted by choosing species or genes that were significantly different in their presence/absence using the Fisher's exact test. Species and genes were sorted separately according to their *P*-value in ascending order and grouped into bins of 100 features. Each bin was used as predictive features for the feed efficiency phenotype using the k-nearest neighbors algorithm (Aha, 1997) with $k = 3$. The mean accuracy of the prediction was calculated using cross-validation of 1000 iterations for each bin, in which 70% of the samples were used as a training set and the remaining 30% were used as a test set to measure the accuracy of the prediction. Changing the bin size (bins ranging in size from 50 to 1000 features per bin) did not affect the accuracy of the prediction (Supplementary Data 3). To check the significance of the classifications accuracy, a permutations technique was used. The classification procedure was repeated 100 times, each time after randomly shuffling (permutating) the sample labels. The *P*-value for each classification accuracy was then obtained by the percentage of permutation runs in which the accuracy achieved was greater than the classification accuracy achieved with the original non-permutated data. The same prediction methodology, accuracy and *P*-value determination were applied to several other metabolic parameters—conversion ratio, milk yield, milk energy, milk lactose, milk fat, milk protein, body conditioning score, pH and DM intake. For each metabolic parameter prediction test, the cows were separated into two groups, by the physiological parameter's mean value (Supplementary Data 4).

For each physiological index, receiver operation characteristics curves and area under curve (AUC) measures were obtained based on the average of 1000 k-nearest neighbors cross-validation iterations. The analysis was performed with the Metrics class that is part of the SKLEARN python machine-learning framework.

Recruitment to microbial genomes and metabolic pathways

Reads from each sample were subsampled according to the sample with lowest number of reads (21 486 100). The reads from each sample were aligned using burrows-wheeler alignment tool to a data set of 59 microbial genomes downloaded from NCBI using burrows-wheeler alignment tool with 98% identity and default parameters. Reads were

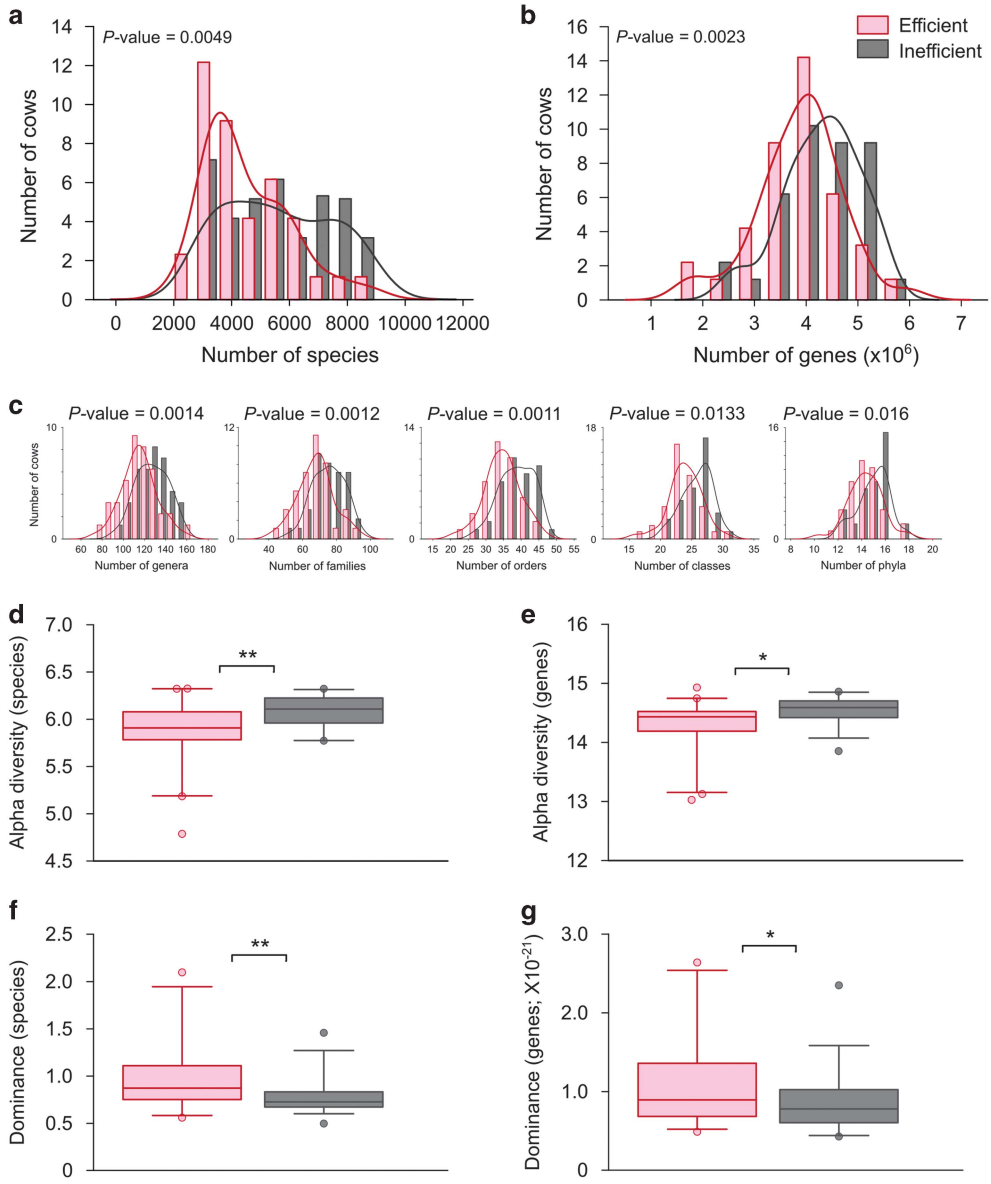

Figure 1 Community parameters of efficient and inefficient cows' microbiomes. (**a**, **b**) Microbiome richness. Species (based on 16S amplicon sequencing) (**a**) and gene (based on metagenomics sequencing) (**b**) counts were calculated and expressed as simple richness. Kernel density of the efficient and inefficient histograms emphasizes the different distribution of counts in each microbiome group. *P*-values of the difference in richness between efficient and inefficient cows are shown. (**c**) Microbiome richness at different phylogenetic levels. (**d**, **e**) Alpha diversity (Shannon index) measurements according to species (**d**) and genes (**e**). (**f**, **g**) Dominance of the microbiome according to species (**f**) and genes (**g**). Data are expressed as mean ± s.e.m. Wilcoxon rank-sum, $*P<0.05$, $**P<0.01$.

also recruited to metabolic pathways of the significantly different metabolites ($P<0.05$) using the same method. Our database consisted of all possible KEGG enzymes for each metabolic pathway. The EC numbers used for each metabolic pathway are described in Supplementary Table S1.

The existence of the propionate production acrylate pathway in the genomes of the examined lactate utilizers *Selenomonas ruminantium* and *Anaerovibrio lipolyticus* was additionally tested by blasting them against all possible KEGG enzymes belonging to the acrylate pathway (EC 1.3.8.7, 2.8.3.1 and 4.2.1.54) using a threshold of above 70% identity, 70% alignment length of the subject gene and $E \leqslant 10^{-5}$.

Results

Construction of a rumen metagenome reference data set
To determine whether there are microbiome features that are associated with the cow's energetic efficiency, the individual feed efficiency of 146 Holstein Friesian cows was first determined. Each animal was automatically monitored for multiple parameters used to calculate feed efficiency (using the RFI approach).

For further analyses, the upper and lower 25% of the animals that exhibited extreme feed efficiency values were chosen, for a total of 78 animals—40 efficient and 38 inefficient (Supplementary Figure S1 and Supplementary Data 1). Metagenomic DNA samples of these animals' rumen microbiomes were subjected to 16S rRNA gene sequencing and whole-genome shotgun sequencing. The metagenomics reads of all samples were pooled and assembled, and the predicted genes served as a reference data set (see Materials and methods section). The metagenome contained 96.72% bacterial sequences, 1.73% archaeal sequences and 1.34% eukaryotic sequences, similar to what was previously described for rumen microbiome metagenomes (Brulc *et al.*, 2009). None of the eukaryotic sequences showed significance in the analyses.

Microbiome features differ and can predict feed efficiency phenotype

A comparison of microbiome richness across the animals revealed significantly lower richness in the efficient cows' microbiomes in both species ($P=0.0049$) and gene content ($P=0.0023$; Figures 1a and b). The differences in taxon richness were apparent up to the phylum level (Figure 1c), further stressing the intensity of this phenomenon. Taxon composition and gene content were derived from two different procedures of sequencing and analysis, and therefore the agreement between these findings highlights the robustness of the observation. The differences in richness were also accompanied by significantly lower diversity and higher dominance in the efficient animals' microbiomes at the species and gene levels ($P<0.01$ and $P<0.05$, respectively, for both diversity and dominance; Figures 1d–g and Supplementary Table S2). These differences were apparent up to the family level (Supplementary Figure S5). These differences in microbiomes of efficient and inefficient cows begged the question of whether microbiome features could be used as markers for the feed efficiency trait.

Thereupon, the species and gene composition of the rumen microbiomes were used to successfully predict the animals' feed efficiency phenotypes with up to 91% accuracy using the k-nearest neighbors algorithm (Aha, 1997). For the feature selection process, a Fisher's exact test was used to measure differences in presence/absence between microbiomes of efficient and inefficient animals. The species and genes were ranked separately according to their *P*-values in ascending order and divided into bins of 100 features to be used for prediction. Each bin was tested for its ability to predict high or low feed efficiency. The mean prediction accuracy was calculated using cross-validation for each bin (1000 iterations). The first species bin's prediction accuracy was 80%, while the first gene bin reached an accuracy of 91% (Figure 2). The species prediction accuracy declined to 50% (accuracy of a random guess) after the fifth bin, whereas the decline in prediction accuracy for the genes followed a much more moderate slope, with the first four predictive bins at above 90% accuracy with highly significant *P*-values. These differences in the slope of prediction accuracy could stem from the fact that each species represents a single genome containing thousands of genes therefore declining more rapidly compared to bins composed of hundreds of single genes. Supplementary Data 5 and 6 contain the identity of species and genes of the first five bins.

The microbiome features were also highly predictive of other physiological parameters, such as milk lactose content and milk yield (Supplementary Figures S6 and S7). The sensitivity and specificity of the predictive bins was further assessed by performing receiver operating characteristic analysis for the first five bins, for both the species and genes data of each physiological parameter (Supplementary Figures S8 and S9). This analysis showed high sensitivity and specificity of the predictions of the host physiological traits based on these microbiome features, as the AUC index had high values that are considered to be good for the species data, and excellent for the genes data

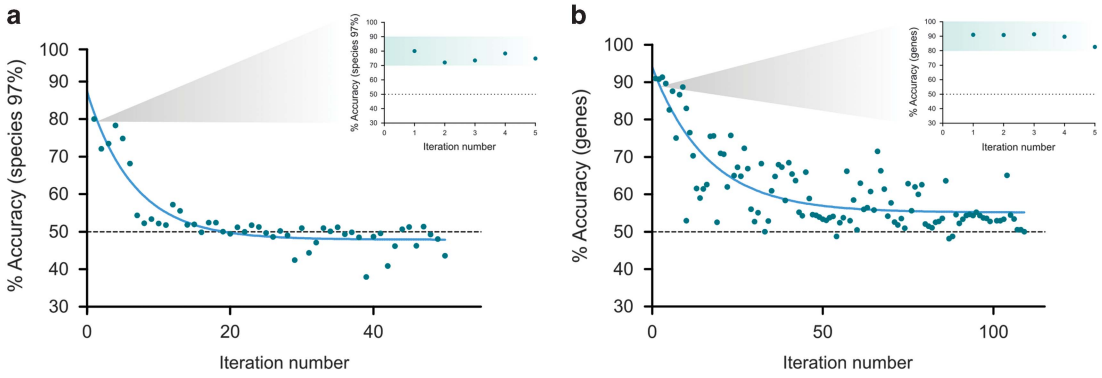

Figure 2 Feed efficiency predictions according to species and genes. Species (**a**) and genes (**b**) that differed in presence/absence between efficient and inefficient cows were ranked according to their *P*-values and grouped into bins of 100. The bins were used as predictive features for the RFI feed efficiency parameter using the k-Nearest Neighbors (KNN) algorithm with $k=3$. Each iteration used a different bin as predictive features, in ascending *P*-value order. Inset in both graphs represents the first five prediction accuracy values (permutations of random classes shuffling, *P*-value = 0.009).

(AUC > 0.8, AUC > 0.9 respectively). This high prediction accuracy indicated that the differences in microbiome gene content and taxonomic composition could be used to classify and predict the cow's energetic efficiency.

Microbiome metabolic activity varies in cows with different feed efficiencies

Diversity, richness and dominance are key ecological determinants that, when altered in a given ecosystem, are expected to have a marked effect on its functionality (Hooper *et al.*, 2005). Hence, following the findings of evident differences in these parameters (Figure 1, Supplementary Figure S5 and Supplementary Table S2), the functionality of the rumen ecosystem was further investigated. Several microbial activity assays, as well as a series of 41 metabolites, were targeted and measured, representing the processes and products of different trophic levels of the rumen microbiome from efficient and inefficient cows, starting from degradation of the ingested plant fiber to the end products (Figure 3).

Significant differences were discovered in most SCFAs. Out of the six SCFAs measured, four—propionate, butyrate, valerate and isovalerate—were at higher concentrations in the rumen of efficient cows (Figure 3, metabolic end products and Supplementary Table S3). In addition, the total concentration of SCFAs was higher in the efficient animals showing an increase of 10% between the two efficiency groups ($P < 0.01$; Figure 4a). These differences are considered to have a marked effect on animal productivity, given that approximately 70% of the net energy requirements of the animal are supplied by SCFAs (Seymour *et al.*, 2005).

Interestingly, the propionate-to-acetate ratio in the efficient animals was also significantly higher than in the inefficient ones ($P < 0.05$; Figure 4b); an increase in this ratio is associated with a decline in methane production and increased energy retention by cattle (Russell, 1998). This finding was congruent with the measurements of the microbiomes' methanogenesis potential, where it was evident that the efficient cows' microbiomes produce significantly less methane than their inefficient counterparts ($P < 0.01$; Figure 3, metabolic end products). The finding of higher concentrations of SCFAs and lower methane emission from the efficient rumen microbiomes is consistent with the notion that

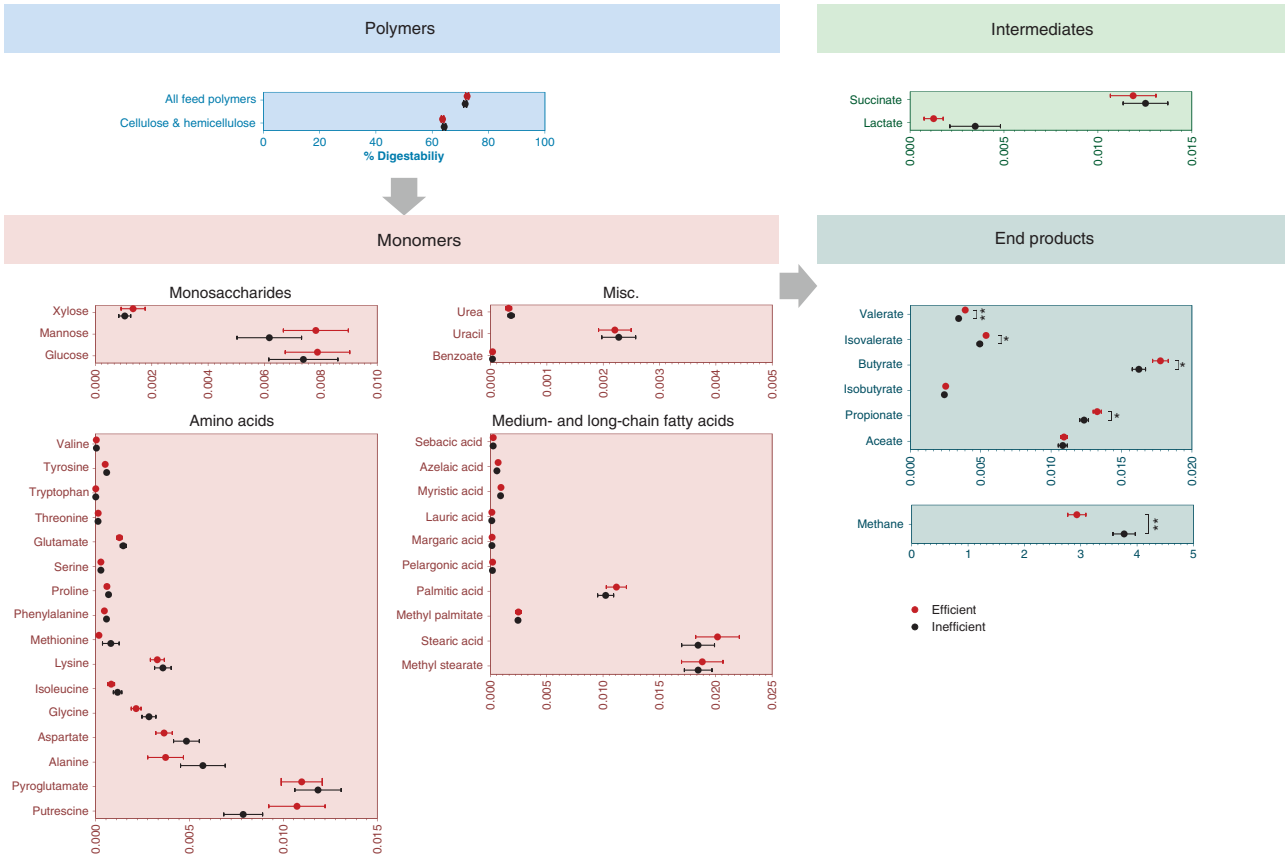

Figure 3 Metabolome and microbial activity of rumen microbiomes of efficient and inefficient cows. *In-vivo* and *in-vitro* digestibility methods were performed on rumen fluid of efficient and inefficient cows in addition to extraction, identification and quantification of 41 different metabolites by GC and gas chromatography mass spectrometry. These metabolites were normalized to the organic matter content of the rumen fluid from which they were extracted. Metabolites are organized according to trophic levels. Multiple hypothesis correction with 9999 permutations was performed individually for each metabolic or activity test using the *t*-statistic (Materials and methods section). Data are expressed as mean ± s.e.m. $^{*}P < 0.05$, $^{**}P < 0.01$.

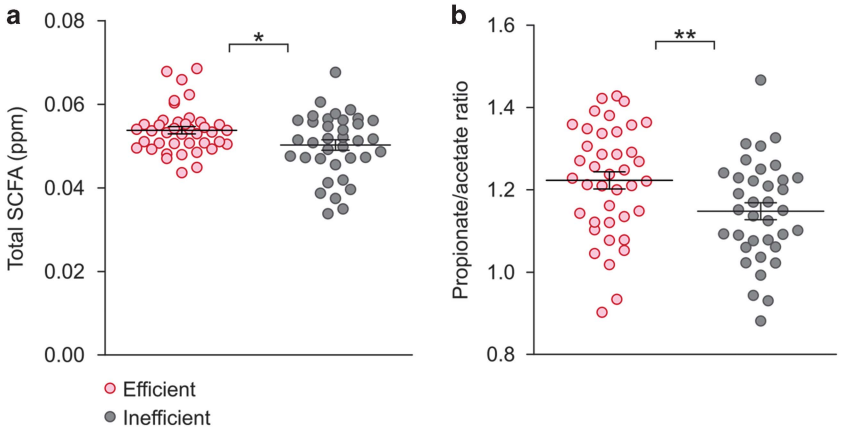

Figure 4 SCFA concentration in rumen fluids of efficient and inefficient cows. (**a**) Total SCFA concentrations in efficient and inefficient rumen samples. (**b**) Propionate/acetate ratio in the efficient and inefficient rumen samples. Data are expressed as mean ± s.e.m. $*P<0.05$, $**P<0.01$.

propionate and butyrate production competes with methanogenesis for hydrogen and presents an alternative mechanism that serves as an electron sink (Ungerfeld, 2015). The production of more SCFAs and less methane by the efficient cows' microbiomes is in agreement with the higher energetic efficiency.

Our analysis did not reveal any significant differences in the microbiomes' ability to degrade the plant cell wall in the diet, *in vitro* or *in vivo* (Figure 3, polymers and Supplementary Figure S10).

Differential abundance of rumen microbes and metabolic pathways

The lower diversity and higher dominance in gene content and taxonomic composition apparent in the microbiomes of efficient cows, together with changes in metabolite assortments, suggested that the flux through collective metabolic pathways is different in microbiomes of this efficiency group. This raised the hypothesis that this may be due to changes in the occupancy of specific rumen microbial niches, defined by metabolic and physical characteristics, by functional groups that differ in their resource demands or output products.

To explore this hypothesis, a permutative Wilcoxon rank-sum test was conducted in which gene and taxonomic profiles were compared between the microbiomes of efficient and inefficient animals (see Materials and methods section). Overall, 18 species and 34 166 genes differentiated the microbiomes of efficient and inefficient cows (Supplementary Figures S11 and S12, and Supplementary Data 7 and 8); of these, 2 species and 227 genes were more abundant in efficient cows. These species and genes were not only differentially abundant in cows with different RFI values, but were also significantly correlated to the intensity of the phenotype (Figure 5a and Supplementary Data 7). The lower numbers of species and genes that were more abundant in the efficient cows' microbiomes are compatible with the higher dominance and lower richness in species and gene composition of these microbiomes. The annotation and analysis of the differentiating genes against the KEGG database (Kanehisa *et al.*, 2011) were also in agreement with these findings, as well as with the metabolomic analysis. Among the KEGG pathways and resultant metabolites that were enriched in the inefficient cows' microbiomes were enzymes from the protein digestion and absorption category, amino-acid biosynthesis and the methane metabolism category (ko numbers of these pathways are detailed in Supplementary Table S4). Furthermore, a significantly lower number of KEGG pathways were enriched in the efficient cows' microbiomes, resulting in a significantly lower number of potential products (Supplementary Figures S13 and S14, Supplementary Data 9).

These findings suggest that there is more diverse use of resource compounds, such as dietary proteins, pyruvate, acetyl-CoA and hydrogen, in the inefficient cows' microbiomes, resulting in a more diverse array of produced metabolites, some of which affect the animal's energy harvest in a negative manner or cannot be utilized by the animal for its energy requirements. In the efficient cows' microbiomes, the use of these compounds is dominated by a limited number of metabolic pathways that are more relevant and valuable for the energy needs of the animal.

The phylogenetic annotations of genes that were enriched in the efficient cows' microbiomes were dominated by the rumen bacterial species *Megasphaera elsdenii*, a highly potent utilizer of lactate for the production of butyrate and propionate (Figure 5b). This annotation, or any other closely related annotation, did not appear in the inefficient cow microbiomes' enriched genes. Overall, the inefficient cows' microbiomes were less dominated by a specific taxon unique to that microbiome group (Figure 5b), further supporting the hypothesis of higher dominance of specific functional groups in the microbiomes of efficient cows. This was also reinforced by the annotation of the two species that were significantly more abundant in the efficient cows' microbiomes in the 16S rRNA gene analysis. One

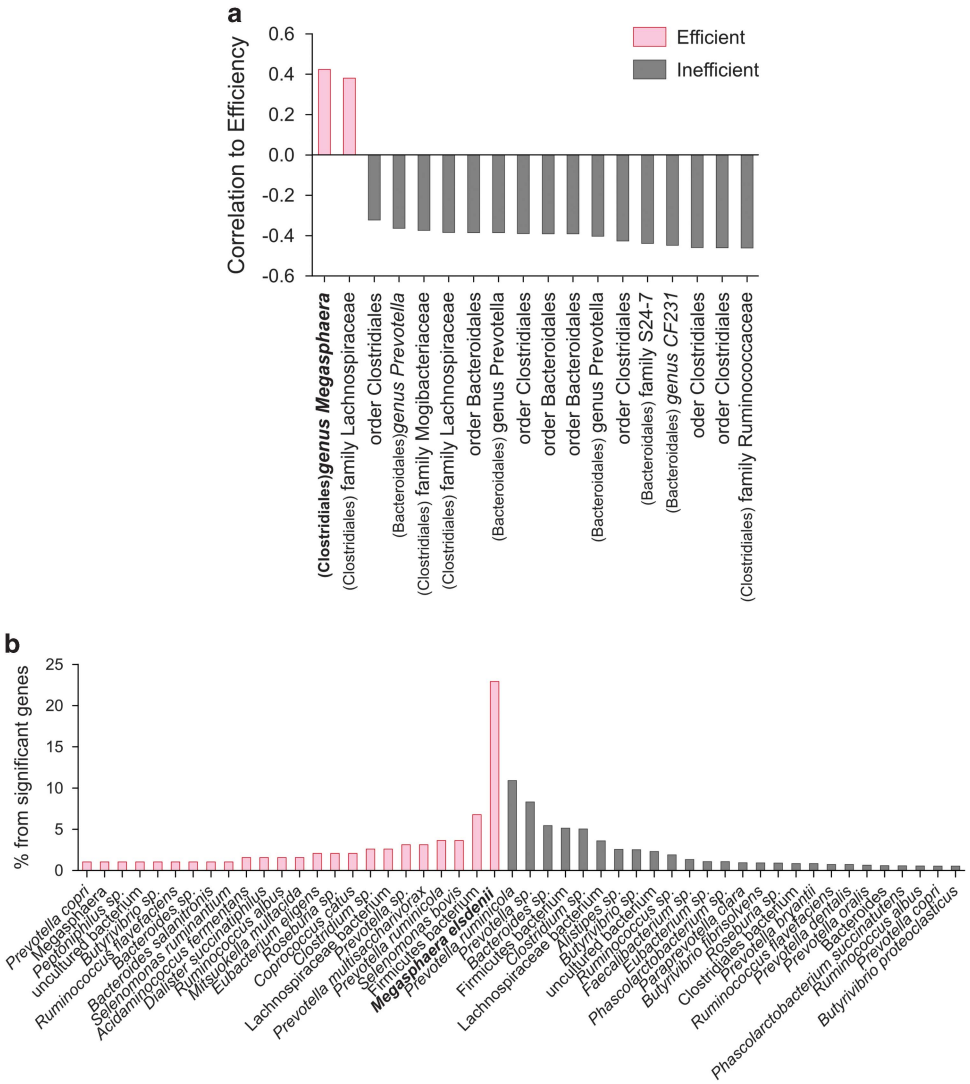

Figure 5 Taxonomic annotations of species and genes enriched in each microbiome group. (**a**) Spearman's correlation of significantly enriched species to the feed efficiency parameter. The annotations are presented at the lowest phylogenetic level obtained, as well as at the order level in parentheses. (**b**) The distribution of the phylogenetic annotations of genes enriched in each of the microbiome groups. Phylogenetic annotations above a threshold of 2% are presented.

annotation that appeared exclusively in this group was of the genus *Megasphaera*. The other abundant species belonged to the family Lachnospiraceae, which also had a representative in the species that were more abundant in the inefficient cows' microbiomes (Figure 5a).

M. elsdenii was also highly enriched in the efficient cows' microbiomes using a different genomic analysis, in which reads from all samples were aligned to a database of 59 sequenced rumen and gut microbial genomes that are known to be involved in various metabolic processes and were also identified in the previous analysis. Here again, inefficient microbiomes were significantly enriched in several microbial genomes, among them *Methanobrevibacter ruminantium* ($P<0.01$), a methanogenic archaeon of the most abundant genus in the rumen (Figure 6a and Supplementary Figure S15). This exploration was further expanded by asking whether these observations are true not only for genomes of specific microbes but for all possible KEGG enzymes belonging to rumen end product metabolic pathways by using the same read-alignment approach (see Materials and methods section). In agreement with the previous results, the methanogenesis pathway was significantly enriched in the inefficient cows' microbiomes ($P<0.01$). Out of all examined pathways for propionate production, only the acrylate pathway that utilizes lactate to propionate was enriched in the efficient cows' microbiomes ($P<0.01$; Figure 6b). It should be noted that this pathway is encoded in the genome of *M. elsdenii* (Prabhu *et al*., 2012) and *Coprococcus catus* (Reichardt *et al*., 2014), which were both found by the analyses to be significantly enriched in efficient animals' microbiomes (Figures 5 and 6a, Supplementary Figure S15), and not in the other examined lactate utilizing microbial genomes (*S. ruminantium* and *A. lipolyticus*). Furthermore, reads

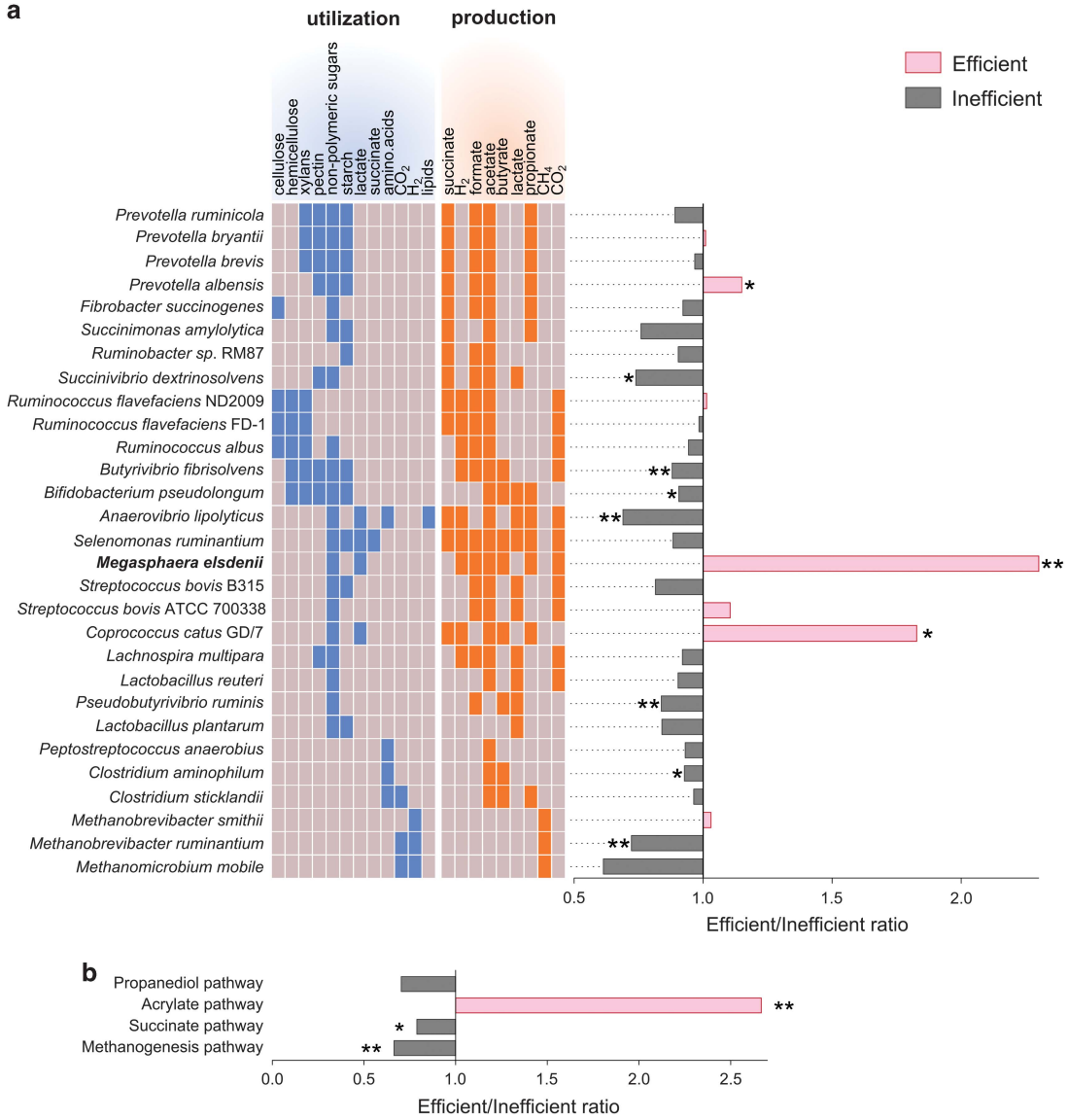

Figure 6 Microbiome features enriched in each microbiome group. (**a**) Reads from each sample were aligned to sequenced genomes of known rumen microorganisms using the burrows-wheeler alignment tool. The ratios between alignments of efficient/inefficient samples to each genome are presented. The utilization and production of metabolites for each microorganism based on the known growth characteristics (Holdman and Moore, 1974; Russell and Rychlik, 2001; Duncan *et al*., 2009) are colored in blue and orange, respectively. (**b**) Reads from each sample were aligned to KEGG enzymes of different metabolic pathways using the burrows-wheeler alignment tool. Propanediol, acrylate and succinate pathways are different propionate production pathways. The ratios between alignments of efficient/inefficient samples to each pathway are presented. Data are expressed as ratio of means. Permutations *t*-test,$^*P<0.05$, $^{**}P<0.01$.

aligned to this pathway are predominantly annotated as *M. elsdenii* and *C. catus*, however, annotations of *Clostridium propionicum* and *Clostridium botulinum* were also detected (Supplementary Figure S16). This highlights the acrylate pathway as the main contributor to the increase in propionate and decrease in lactate observed in the metabolomic analysis of the efficient cows' microbiome group (Figure 3).

Discussion

Our analyses of multiple animals feeding on the exact same diet and kept under the same conditions showed that there are large variations in the individual animals' ability to extract energy from their feed. These variations are tightly linked to several microbiome features that include a decrease in richness and increase in dominance of taxonomic and coding capacity in the efficient cow's microbiome. They are reflected as changes in this ecosystem's functionality, where changes in the dominance of specific functional components affect the overall availability of ecosystem goods that are of high value to the hosting animal. Higher microbiome richness and changes in specific functional groups have been recently described to affect host productivity in plants (Wagg *et al*., 2014) as well as humans, where lower diversity

and richness have been associated with higher energy harvest from feed in obese humans (Turnbaugh *et al.*, 2009; Le Chatelier *et al.*, 2013). A possible explanation for this phenomenon could stem from a more diverse use of resource compounds in the inefficient cow's microbiomes that are enriched in species, genes and KEGG pathways resulting in a wider array of output metabolites (Figures 3, 6 and 7); this was also confirmed by significantly more KEGG output metabolites (Supplementary Figures S13 and S14). On the other hand, in the efficient cow's microbiome, simpler metabolic pathway networks result in increased dominance of specific functional components, which leads to higher concentrations of ecosystem goods that are relevant to the host (Figure 7b). Therefore, the efficient microbiomes are less complex but more specialized to support the host's energy requirements.

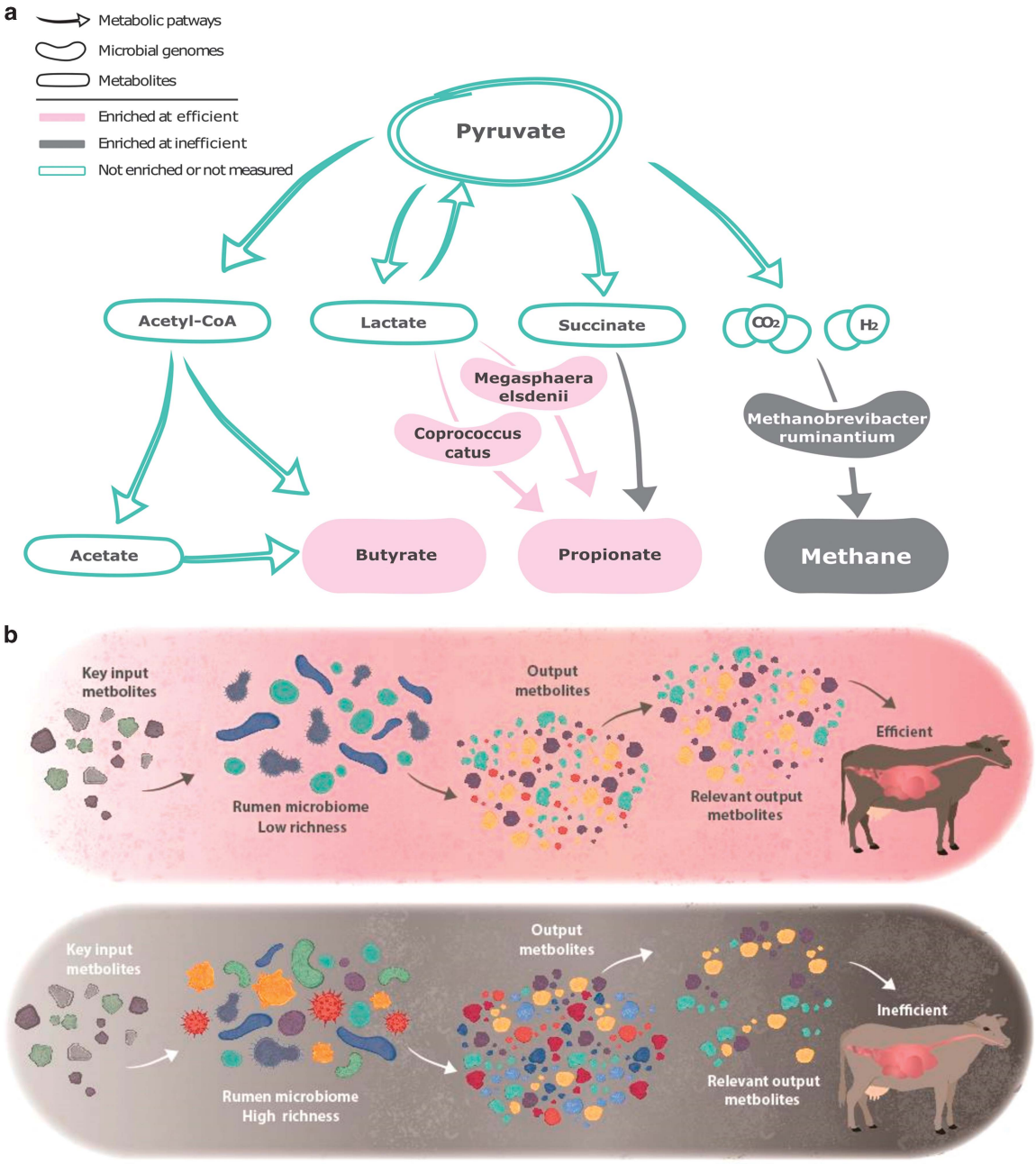

Figure 7 Consolidated results and model. (**a**) Consolidation of results from the metabolomics, genome and pathway recruitment analyses. Green: pathways and metabolites that were not significantly different or that were not assessed. Pink: enriched in efficient microbiomes. Grey: enriched in inefficient microbiomes. (**b**) Proposed model. From left to right: identical key input metabolites are ingested by the cow and presented to either an efficient microbiome (top panel) with lower richness and diversity, or an inefficient microbiome (bottom panel) with higher richness and diversity. Differences in richness result in the production of different metabolites. The efficient microbiome produces a smaller range of output metabolites than the inefficient microbiome, however, with larger amounts of relevant output metabolites, which are available for the animal's energetic needs.

This notion is exemplified by the finding of higher concentrations of SCFAs, which are valuable to the hosting animal. SCFAs are absorbed through the rumen wall to serve the energetic needs of the animal; propionate, for example, is the main precursor for gluconeogenesis in animals (Russell and Wilson, 1996; Mizrahi, 2011; Mizrahi, 2013). This is not the case with methane as the energy retained in it cannot be absorbed by the animals, and is lost to the atmosphere. Such metabolic changes are usually achieved via the use of antibiotic growth promoters that increase the animal's feed efficiency (Duffield *et al*., 2012). Such is the case with monensin, a carboxylic polyether ionophore that selectively affects some of the rumen microbes, therefore changing the structure of the rumen microbiome and subsequently the ratio of SCFAs in the rumen, increasing propionic acid and decreasing methane production (Thornton and Owens, 1981; Callaway *et al*., 2003; Weimer *et al*., 2008; Duffield *et al*., 2012). It has been shown that when administered orally, monensin improves feed efficiency in cattle in a dose-dependent manner. Therefore, it has been used for this purpose extensively since its approval for cattle agriculture in the mid-1970s (Duffield *et al*., 2012). This effect of rumen microbiome manipulation achieved via antibiotics further supports the connection of the rumen microbiome with the feed efficiency of the animal.

Here we show that these metabolomic changes are the outcome of microbiome structures that are naturally occurring and are highly correlated with, and predictive of the feed efficiency phenotype. Therefore, these findings could be harnessed to reduce the use of antibiotic growth promoters in agriculture. Such prospects, together with examination of the true causal nature of the rumen microbiome on its host, would be made possible when non-antibiotic rumen microbiome manipulation techniques, as well as adult cow germ-free facilities, will be better established.

From an ecological perspective, the lower abundance of methanogenesis pathways and methanogenic archaea in the efficient cow's low-richness microbiome concurs with the notion that processes that are performed by small taxonomic groups, such as the methanogenic archaea that occupy only small percentages of the rumen microbiome, are more sensitive to changes in diversity and richness (Hooper *et al*., 1995). These changes are usually accompanied by occupation and dominance of the available niche by different species using the same resources (Grime, 1998). Such is the case with *M. elsdenii* and *C. catus*, independently found to be enriched in the efficient animals' microbiomes in different analyses (Figures 5 and 6a, Supplementary Figure S15), which use electrons for the production of the valuable SCFAs propionate and butyrate, thereby diverting them from reducing CO_2 to methane (Prabhu *et al*., 2012; Ungerfeld, 2015). A similar principle was shown to apply in Tammar wallabies, where *Succinivibrio* bacteria were suggested to utilize hydrogen for the production of succinate, therefore lowering its availability for methanogenesis (Pope *et al*., 2011). It is also possible that the Lachnospiraceae detected in the efficient animals' microbiomes are butyrate producers (Figure 5) and are contributing further to this effect (Louis and Flint, 2009; Meehan and Beiko, 2014). Nevertheless, as other SCFAs are enriched in this microbiome group and most of the carbon flux in the system goes to acetyl-CoA, formate or hydrogen and carbon dioxide, it is likely that more genes and pathways are involved in this effect.

A cardinal point that emerges from our findings is that the functional characteristics of a small number of species can have a large impact on community structure and ecosystem functioning. This, in turn, can change the productivity of the supraorganism—the host and its residing rumen microbiome.

These findings could potentially be harnessed to increase the production of food resources for mankind in a more sustainable manner, as well as to understand the underlying ecological mechanisms that govern complex microbial communities and their interactions with their hosts.

Conflict of Interest

The authors declare no conflict of interest.

Acknowledgements

This research was supported by Binational Agricultural Research and Development fund project number 362032713, the Israel Science Foundation project number 1313/13 and the European Research Council under the European Union's Horizon 2020 research and innovation program, project number 640384.

Data deposition

Assembled Contigs were deposited in the MG-RAST server under IDs 4547164.3 and 4548996.3.

References

Aha DW. (1997). *Lazy Learning*. Kluwer Academic Publishers: Norwell, MA, USA.

Ajmone-Marsan P, Garcia JF, Lenstra JA. (2010). On the origin of cattle: how aurochs became cattle and colonized the world. *Evol Anthropol Issues News Rev* **19**: 148–157.

Archer JA, Richardson EC, Herd RM, Arthur PF. (1999). Potential for selection to improve efficiency of feed use in beef cattle. *Aust J Agric Res* **50**: 147–162.

Benjamini Y, Hochberg Y. (1995). Controlling the false discovery rate: a practical and powerful approach to multiple testing. *J R Stat Soc Series B (Methodological)* **57**: 289–300.

Bradford GE. (1999). Contributions of animal agriculture to meeting global human food demand. *Livest Prod Sci* **59**: 95–112.

Brulc JM, Antonopoulos DA, Miller ME, Wilson MK, Yannarell AC, Dinsdale EA *et al.* (2009). Gene-centric metagenomics of the fiber-adherent bovine rumen microbiome reveals forage specific glycoside hydrolases. *Proc Natl Acad Sci USA* **106**: 1948–1953.

Callaway TR, Edrington TS, Rychlik JL, Genovese KJ, Poole TL, Jung YS *et al.* (2003). Ionophores: their use as ruminant growth promotants and impact on food safety. *Curr Issues Intest Microbiol* **4**: 43–51.

Caporaso JG, Kuczynski J, Stombaugh J, Bittinger K, Bushman FD, Costello EK *et al.* (2010). QIIME allows analysis of high-throughput community sequencing data. *Nat Methods* **7**: 335–336.

Cole JR, Chai B, Marsh TL, Farris RJ, Wang Q, Kulam SA *et al.* (2003). The Ribosomal Database Project (RDP-II): previewing a new autoaligner that allows regular updates and the new prokaryotic taxonomy. *Nucleic Acids Res* **31**: 442–443.

Davis JC. (1986). *Statistics and Data Analysis in Geology*. Wiley: New York.

Duffield TF, Merrill JK, Bagg RN. (2012). Meta-analysis of the effects of monensin in beef cattle on feed efficiency, body weight gain, and dry matter intake. *J Anim Sci* **90**: 4583–4592.

Duncan SH, Louis P, Thomson JM, Flint HJ. (2009). The role of pH in determining the species composition of the human colonic microbiota. *Environ Microbiol* **11**: 2112–2122.

Edgar RC. (2010). Search and clustering orders of magnitude faster than BLAST. *Bioinformatics* **26**: 2460–2461.

Grime J. (1998). Benefits of plant diversity to ecosystems: immediate, filter and founder effects. *J Ecol* **86**: 902–910.

Halachmi I, Edan Y, Maltz E, Peiper U, Moallem U, Brukental I. (1998). A real-time control system for individual dairy cow food intake. *Comput Electron Agr* **20**: 131–144.

Harper D. (1999). *Numerical Palaeobiology*. Wiley: London.

Henderson G, Cox F, Ganesh S, Jonker A, Young W, Collaborators GRC *et al.* (2015). Rumen microbial community composition varies with diet and host, but a core microbiome is found across a wide geographical range. *Sci Rep* **5**: 14567.

Herd RM, Arthur PF. (2009). Physiological basis for residual feed intake. *J Anim Sci* **87**: E64–E71.

Hernandez-Sanabria E, Goonewardene LA, Wang Z, Durunna ON, Moore SS. (2012). Impact of feed efficiency and diet on adaptive variations in the bacterial community in the rumen fluid of cattle. *Appl Environ Microb* **78**: 1203–1214.

Hess M, Sczyrba A, Egan R, Kim TW, Chokhawala H, Schroth G *et al.* (2011). Metagenomic discovery of biomass-degrading genes and genomes from cow rumen. *Science* **331**: 463–467.

Holdman LV, Moore WEC. (1974). New genus, *Coprococcus*, twelve new species, and emended descriptions of four previously described species of bacteria from human feces. *Int J Syst Evol Micr* **24**: 260–277.

Hooper DU, Hawksworth D, Dhillion S. (1995). Microbial diversity and ecosystem processes. Global biodiversity assessment*UNEP*. Cambridge University Press: Cambridge, pp 433–443.

Hooper DU, Chapin FS, Ewel JJ, Hector A, Inchausti P, Lavorel S *et al.* (2005). Effects of biodiversity on ecosystem functioning: a consensus of current knowledge. *Ecol Monogr* **75**: 3–35.

Jami E, Israel A, Kotser A, Mizrahi I. (2013). Exploring the bovine rumen bacterial community from birth to adulthood. *ISME J* **7**: 1069–1079.

Jami E, Mizrahi I. (2012). Composition and similarity of bovine rumen microbiota across individual animals. *PLoS One* **7**: e33306.

Jami E, White BA, Mizrahi I. (2014). Potential role of the bovine rumen microbiome in modulating milk composition and feed efficiency. *PLoS One* **9**: e85423.

Pollard KS, Dudoit S, van der Laan MJ. (2005). Multiple testing procedures: the multtest package and applications to genomics. In: Gentleman R, Carey V, Huber W, Irizarry R, Dudoit S (eds), *Bioinformatics and Computational Biology Solutions Using R and Bioconductor*. Springer: New York, USA, pp 249–271.

Kanehisa M, Goto S, Sato Y, Furumichi M, Tanabe M. (2011). KEGG for integration and interpretation of large-scale molecular data sets. *Nucleic Acids Res* **40**: D109–D114.

Kittelmann S, Pinares-Patiño CS, Seedorf H, Kirk MR, Ganesh S, Mcewan JC *et al.* (2014). Two different bacterial community types are linked with the low-methane emission trait in sheep. *PLoS One* **9**: e103171.

Koch RM, Swiger LA, Chambers D, Gregory KE. (1963). Efficiency of feed use in beef cattle. *J Anim Sci* **22**: 486–494.

Kultima JR, Sunagawa S, Li J, Chen W, Chen H, Mende DR *et al.* (2012). MOCAT: a metagenomics assembly and gene prediction toolkit. *PLoS One* **7**: e47656.

Le Chatelier E, Nielsen T, Qin J, Prifti E, Hildebrand F, Falony G *et al.* (2013). Richness of human gut microbiome correlates with metabolic markers. *Nature* **500**: 541–546.

Li D, Lewinger JP, Gauderman WJ, Murcray CE, Conti D. (2011). Using extreme phenotype sampling to identify the rare causal variants of quantitative traits in association studies. *Genet Epidemiol* **35**: 790–799.

Li H, Durbin R. (2009). Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics* **25**: 1754–1760.

Lippke H, Ellis WC, Jacobs BF. (1986). Recovery of indigestible fiber from feces of sheep and cattle on forage diets. *J Dairy Sci* **69**: 403–412.

Louis P, Flint HJ. (2009). Diversity, metabolism and microbial ecology of butyrate-producing bacteria from the human large intestine. *FEMS Microbiol Lett* **294**: 1–8.

McDougall E. (1948). Studies on ruminant saliva. 1. The composition and output of sheep's saliva. *Biochem J* **43**: 99.

McMichael AJ, Powles JW, Butler CD, Uauy R. (2007). Food, livestock production, energy, climate change, and health. *Lancet* **370**: 1253–1263.

Meehan CJ, Beiko RG. (2014). A phylogenomic view of ecological specialization in the Lachnospiraceae, a family of digestive tract-associated bacteria. *Genome Biol Evol* **6**: 703–713.

Meyer F, Paarmann D, D'souza M, Olson R, Glass E, Kubal M *et al.* (2008). The metagenomics RAST server—a public resource for the automatic phylogenetic and functional analysis of metagenomes. *BMC Bioinform* **9**: 386.

Mizrahi I. (2011). The role of the Rumen microbiota in determining the feed efficiency of dairy cows. In: Rosenberg E, Gophna U (eds), *Beneficial Microorganisms in Multicellular Life Forms*. Springer Science & Business Media: Berlin Heidelberg, pp 203–210.

Mizrahi I. (2013). Rumen symbioses. In: Rosenberg E, DeLong F, Lory S, Stackebrandt E, Thompson F (eds), *The Prokaryotes*. Springer Science & Business Media: Berlin Heidelberg, pp 533–544.

Moore SS, Mujibi FD, Sherman EL. (2009). Molecular basis for residual feed intake in beef cattle. *J Anim Sci* **87**: E41–E47.

National Research Council. (2001). *Nutrient Requirements of Dairy Cattle*. National Academy Press Washington: DC, USA.

Nkrumah JD, Okine EK, Mathison GW, Schmid K, Li C, Basarab JA *et al.* (2006). Relationships of feedlot feed efficiency, performance, and feeding behavior with metabolic rate, methane production, and energy partitioning in beef cattle. *J Anim Sci* **84**: 145–153.

Peterson J, Garges S, Giovanni M, Mcinnes P, Wang L, Schloss JA *et al.* (2009). The NIH human microbiome project. *Genome Res* **19**: 2317–2323.

Pope PB, Smith W, Denman SE, Tringe SG, Barry K, Hugenholtz P *et al.* (2011). Isolation of Succinivibrionaceae implicated in low methane emissions from Tammar wallabies. *Science* **333**: 646–648.

Prabhu R, Altman E, Eiteman MA. (2012). Lactate and acrylate metabolism by *Megasphaera elsdenii* under batch and steady-state conditions. *Appl Environ Microbiol* **78**: 8564–8570.

Pruitt KD, Tatusova T, Maglott DR. (2007). NCBI reference sequences (RefSeq): a curated non-redundant sequence database of genomes, transcripts and proteins. *Nucleic Acids Res* **35**: D61–D65.

Pryce J, Arias J, Bowman P, Davis S, Macdonald K, Waghorn G *et al.* (2012). Accuracy of genomic predictions of residual feed intake and 250-day body weight in growing heifers using 625,000 single nucleotide polymorphism markers. *J Dairy Sci* **95**: 2108–2119.

Reichardt N, Duncan SH, Young P, Belenguer A, Mcwilliam Leitch C, Scott KP *et al.* (2014). Phylogenetic distribution of three pathways for propionate production within the human gut microbiota. *ISME J* **8**: 1323–1335.

Rho M, Tang H, Ye Y. (2010). FragGeneScan: predicting genes in short and error-prone reads. *Nucleic Acids Res* **38**: e191.

Russell JB. (1998). The importance of pH in the regulation of ruminal acetate to propionate ratio and methane production *in vitro*. *J Dairy Sci* **81**: 3222–3230.

Russell JB, Rychlik JL. (2001). Factors that alter rumen microbial ecology. *Science* **292**: 1119–1122.

Russell JB, Wilson DB. (1996). Why are ruminal cellulolytic bacteria unable to digest cellulose at low pH? *J Dairy Sci* **79**: 1503–1509.

Saleem F, Bouatra S, Guo AC, Psychogios N, Mandal R, Dunn SM *et al.* (2013). The bovine ruminal fluid metabolome. *Metabolomics* **9**: 360–378.

Seymour WM, Campbell DR, Johnson ZB. (2005). Relationships between rumen volatile fatty acid concentrations and milk production in dairy cows: a literature study. *Anim Feed Sci Tech* **119**: 155–169.

Shi W, Moon CD, Leahy SC, Kang D, Froula J, Kittelmann S *et al.* (2014). Methane yield phenotypes linked to differential gene expression in the sheep rumen microbiome. *Genome Res* **24**: 1517–1525.

Stevenson DM, Weimer PJ. (2007). Dominance of Prevotella and low abundance of classical ruminal bacterial species in the bovine rumen revealed by relative quantification real-time PCR. *Appl Microbiol Biotechnol* **75**: 165–174.

Thornton JH, Owens FN. (1981). Monensin supplementation and *in vivo* methane production by steers. *J Anim Sci* **52**: 628–634.

Thornton PK. (2010). Livestock production: recent trends, future prospects. *Philos Trans R Soc Lond B Biol Sci* **365**: 2853–2867.

Tilley J, Terry R. (1963). A two-stage technique for the in vitro digestion of forage crops. *Grass Forage Sci* **18**: 104–111.

Turnbaugh PJ, Hamady M, Yatsunenko T, Cantarel BL, Duncan A, Ley RE *et al.* (2009). A core gut microbiome in obese and lean twins. *Nature* **457**: 480–484.

Ungerfeld EM. (2015). Shifts in metabolic hydrogen sinks in the methanogenesis-inhibited ruminal fermentation: a meta-analysis. *Front Microbiol* **6**: 37.

Van Soest PJ, Robertson JB, Lewis BA. (1991). Methods for dietary fiber, neutral detergent fiber, and nonstarch polysaccharides in relation to animal nutrition. *J Dairy Sci* **74**: 3583–3597.

Wagg C, Bender SF, Widmer F, van der Heijden MG. (2014). Soil biodiversity and soil community composition determine ecosystem multifunctionality. *Proc Natl Acad Sci USA* **111**: 5266–5270.

Wallace RJ, Rooke JA, Mckain N, Duthie CA, Hyslop JJ, Ross DW *et al.* (2015). The rumen microbial metagenome associated with high methane production in cattle. *BMC Genomics* **16**: 1.

Weimer PJ, Stevenson DM, Mertens DR, Thomas EE. (2008). Effect of monensin feeding and withdrawal on populations of individual bacterial species in the rumen of lactating dairy cows fed high-starch rations. *Appl Microbiol Biotechnol* **80**: 135–145.

Westfall PH, Young SS. (1993). *Resampling-Based Multiple Testing: Examples and Methods for P-Value Adjustment*. John Wiley & Sons: New York.

This work is licensed under a Creative Commons Attribution-NonCommercial-ShareAlike 4.0 International License. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in the credit line; if the material is not included under the Creative Commons license, users will need to obtain permission from the license holder to reproduce the material. To view a copy of this license, visit http://creativecommons.org/licenses/by-nc-sa/4.0/

Supplementary Information accompanies this paper on The ISME Journal website (http://www.nature.com/ismej)

AMERICAN SOCIETY FOR MICROBIOLOGY | mBio®

RESEARCH ARTICLE

Heritable Bovine Rumen Bacteria Are Phylogenetically Related and Correlated with the Cow's Capacity To Harvest Energy from Its Feed

Goor Sasson,[a] Sheerli Kruger Ben-Shabat,[a,b,c] Eyal Seroussi,[b] Adi Doron-Faigenboim,[b] Naama Shterzer,[a] Shamay Yaacoby,[b] Margret E. Berg Miller,[d,e] Bryan A. White,[d,e] Eran Halperin,[c,f] Itzhak Mizrahi[a]

Department of Life Sciences and the National Institute for Biotechnology in the Negev, Ben-Gurion University of the Negev, Beer-Sheva, Israel[a]; Department of Ruminant Science, Institute of Animal Sciences, Agricultural Research Organization, Volcani Center, Bet Dagan, Israel[b]; Department of Molecular Microbiology and Biotechnology, The George S. Wise Faculty of Life Science, Tel Aviv University, Ramat-Aviv, Israel[c]; Department of Animal Sciences, University of Illinois, Urbana, Illinois, USA[d]; The Institute for Genomic Biology, University of Illinois, Urbana, Illinois, USA[e]; Blavatnik School of Computer Science, Tel Aviv University, Tel Aviv, Israel[f]

ABSTRACT Ruminants sustain a long-lasting obligatory relationship with their rumen microbiome dating back 50 million years. In this unique host-microbiome relationship, the host's ability to digest its feed is completely dependent on its coevolved microbiome. This extraordinary alliance raises questions regarding the dependent relationship between ruminants' genetics and physiology and the rumen microbiome structure, composition, and metabolism. To elucidate this relationship, we examined the association of host genetics with the phylogenetic and functional composition of the rumen microbiome. We accomplished this by studying a population of 78 Holstein-Friesian dairy cows, using a combination of rumen microbiota data and other phenotypes from each animal with genotypic data from a subset of 47 animals. We identified 22 operational taxonomic units (OTUs) whose abundances were associated with rumen metabolic traits and host physiological traits and which showed measurable heritability. The abundance patterns of these microbes can explain high proportions of variance in rumen metabolism and many of the host physiological attributes such as its energy-harvesting efficiency. Interestingly, these OTUs shared higher phylogenetic similarity between themselves than expected by chance, suggesting occupation of a specific ecological niche within the rumen ecosystem. The findings presented here suggest that ruminant genetics and physiology are correlated with microbiome structure and that host genetics may shape the microbiome landscape by enriching for phylogenetically related taxa that may occupy a unique niche.

IMPORTANCE Dairy cows are an essential nutritional source for the world's population; as such, they are extensively farmed throughout our planet and subsequently impact our environment. The microbial communities that reside in the upper digestive tract of these animals in a compartment named the rumen degrade and ferment the plant biomass that the animal ingests. Our recent efforts, as well as those of others, have shown that this microbial community's composition and functionality are tightly linked to the cow's capacity to harvest energy from its feed, as well as to other physiological traits. In this study, we identified microbial groups that are heritable and also linked to the cow's production parameters. This finding could potentially allow us to apply selection programs on specific rumen microbial components that are linked to the animal's physiology and beneficial to production. Hence, it is a

Received 22 May 2017 **Accepted** 1 June 2017 **Published** 15 August 2017

Citation Sasson G, Kruger Ben-Shabat S, Seroussi E, Doron-Faigenboim A, Shterzer N, Yaacoby S, Berg Miller ME, White BA, Halperin E, Mizrahi I. 2017. Heritable bovine rumen bacteria are phylogenetically related and correlated with the cow's capacity to harvest energy from its feed. mBio 8:e00703-17. https://doi.org/10.1128/mBio.00703-17.

Editor Margaret J. McFall-Ngai, University of Hawaii at Manoa

Copyright © 2017 Sasson et al. This is an open-access article distributed under the terms of the Creative Commons Attribution 4.0 International license.

Address correspondence to Itzhak Mizrahi, imizrahi@bgu.ac.il.

This article is a direct contribution from a Fellow of the American Academy of Microbiology. Solicited external reviewers: Andrew Benson, University of Nebraska–Lincoln; Dan Knights, University of Minnesota.

Downloaded from http://mbio.asm.org/ on June 2, 2019 by guest

steppingstone toward microbiome manipulation for increasing food availability while lowering environmental impacts such as methane emission.

KEYWORDS genetics, host-microbe interaction, microbial ecology, microbiome, rumen ecology

The bovine rumen microbiome essentially enables the hosting ruminant animal to digest its feed by degrading and fermenting it. In this sense, this relationship is unique and different from the host-microbiome interactions that have evolved in humans and nonherbivorous animals, where such dependence does not exist (1, 2). This strict obligatory host-microbiome relationship, which was established approximately 50 million years ago, is thought to play a major role in host physiology (3). Despite its great importance, the impact of natural genetic variation in the host—brought about through sexual reproduction and meiotic recombination—on the complex relationship of rumen microbiome components and host physiological traits is poorly understood. Indeed, several works have reported such a link with regard to methane emission (4–7) and energy-harvesting efficiency (8–12) and, recently, associations between specific components of the rumen microbiome and animal physiology, mainly exemplified by the ability of the animal to harvest energy from its feed (13). These recent findings position the bovine rumen microbiome as the new frontier in the effort to increase the feed efficiency of dairy cows. As the human population is continually increasing, this could have important implications for food security issues as an effort toward replenishing food sources available for human consumption while lowering environmental impact on a global scale. Despite its great importance, the complex relationship of rumen microbiome components and host genetics and physiology is poorly understood. While a well-established connection exists between the structure and function of the rumen microbiome and dietary regime (14–18) as well as specific physical rumen traits such as redox potential (19, 20), until now only three studies have addressed the question of host genetics' interaction with the rumen microbiome. In one study, PCR-denaturing gradient gel electrophoresis (DGGE) profiles of rumen microbiome samples from 18 steers of different breeds varying in their feed efficiency were compared (11). Although no direct correlation was found, some of the animals clustered as a function of their breed; therefore, it was suggested that host genetics may play a role in rumen microbiome structure. Recently, support for the interaction of host genetics with rumen microbiome composition came from a comparison of archaeon/bacterium ratios in the rumen microbiomes of eight animals from different breeds. In this study, the animals' progeny groups were correlated with methane emissions and with archaeon/bacterium ratio (21), possibly suggesting that host genetics is connected to the ratio between these two domains. A later research study on deer hybrids showed that hybrid offspring have different microbial compositions than their parents. Additionally, alanine, arginine, proline, and phenylalanine pathways were enriched in hybrid offspring, and this enrichment was correlated with the abundance of the bacterial spp. *Prevotella*, *Acetitomaculum*, *Quinella*, *Succinivibrio*, and *Ruminobacter* (22).

Here, we explored the interaction of host genetics with bovine rumen microbiome components, with the aim of identifying specific microbes whose abundances are influenced by genetic variation in the host. We further aimed at exploring whether such identified components could be connected to rumen metabolism and host physiological attributes. To this end, we used bovine single nucleotide polymorphism chips (SNP chips) as well as bovine microbial taxa that were inferred from 16 rRNA amplicon sequencing data. We first determined the genetic relatedness between cows based on genomic SNP similarity between the animals. We then combined that information with the abundance profiles of microbiome components across animals and estimated their heritability. Our study addressed the question of whether species in the bovine rumen are heritable and the taxonomic and phylogenetic relationships of species that show measurable heritability.

Downloaded from http://mbio.asm.org/ on June 2, 2019 by guest

Finally, we asked whether the heritable microbial species may be associated with important host physiological traits and metabolic traits of the rumen.

RESULTS

Experimental design and data. Our main goal was to identify microbial species where significant proportions of their variation in abundance profiles can be attributed to heritable genetic factors. To achieve this, we analyzed common SNP genotyping information of 47 Holstein-Friesian dairy cows. We consolidated this information with additional data for these animals from our recent study (13) in which we have discovered microbiome mechanisms that underlie energy-harvesting efficiency from feed in bovines. For each animal, we sequenced the 16S rRNA gene from rumen samples from three consecutive days. We also quantified rumen metabolites and performed rumen metabolic activity assays such as *ex vivo* rumen methane production and fiber digestion measurements. We also consolidated metadata of individual cows' production indices and physiological indices. Low-quality and noninformative SNPs were removed using a quality control (QC) pipeline (see Materials and Methods). 16S amplicon sequencing analysis was performed using the QIIME pipeline (23). The rumen microbiome taxonomic profiles represented by 85,255 species-level operational taxonomic units (OTUs) (three samples per animal) were associated with genomic data represented by genotyping of common SNP loci (see Materials and Methods). Notably, we focused on identification of heritable microbial OTUs rather than the heritability estimate magnitude. This approach is more robust to heritability estimate values that are typical with small sample sizes in estimation procedures. The microbial OTUs found to be associated with the animals' genomes were further correlated with metabolomics data of the microbiomes, as well as with animal physiology and productivity parameters.

Heritable species have high phylogenetic relatedness and are enriched with the order *Bacteroidales*. The first step in our analysis was to identify heritable bacterial species; that is, in our context, to identify microbial species where significant proportions of their variation in abundance profiles could be attributed to heritable genetic factors. This would be reflected in a highly similar abundance of certain species among animals that share a similar genetic background. Accordingly, we estimated the relatedness between all pairs of animals in the cohort. This estimation was done by considering both the count and the infrequency of the alleles (SNPs) in the reference genotypes. We used these pairwise genetic relationship estimates together with each species' abundance profile to calculate their heritability estimate (Fig. 1).

To increase the confidence of our analysis, we limited our heritability analysis exclusively to OTUs which were present in at least 12 genotyped animals (25% of the genotyped subset) as previously described (24). In addition, we performed three independent heritability analyses for each OTU, one for each sampling day. Only OTUs that exhibited a significant heritable component (heritability estimate of >0.7 and P value of <0.05) in all three individual sampling days were considered heritable OTUs. Following this procedure, our analysis resulted in 22 heritable OTUs that match these criteria (see Fig. S3 in the supplemental material), all belonging to the bacterial domain. Although the heritability significance assessment procedure is based on a parametric test, we inspected the robustness of this finding by examining the false discovery rate distribution of the test under permuted assumptions. For that purpose, we generated a null model with 100 iterations, where in each iteration we repeated the heritability analyses after randomly shuffling the genetic profile order. In 94% of the permutations, the number of OTUs detected as heritable was smaller than 22, while in most permutations, the number of OTUs detected as heritable was under 5 (Fig. S1).

It is interesting that the heritable OTUs exhibited a high presence across animals, ranging between 50% and 100% of the animals, with the majority appearing in 70 to 100% of the examined animals (Fig. 2; Fig. S4 and S5). The abundance profile of the heritable microbes was correlated with their presence profile (Spearman correlation between the presence counts and abundance sums, $r = 0.75$, $P < 5 \times 10^{-5}$).

Downloaded from http://mbio.asm.org/ on June 2, 2019 by guest

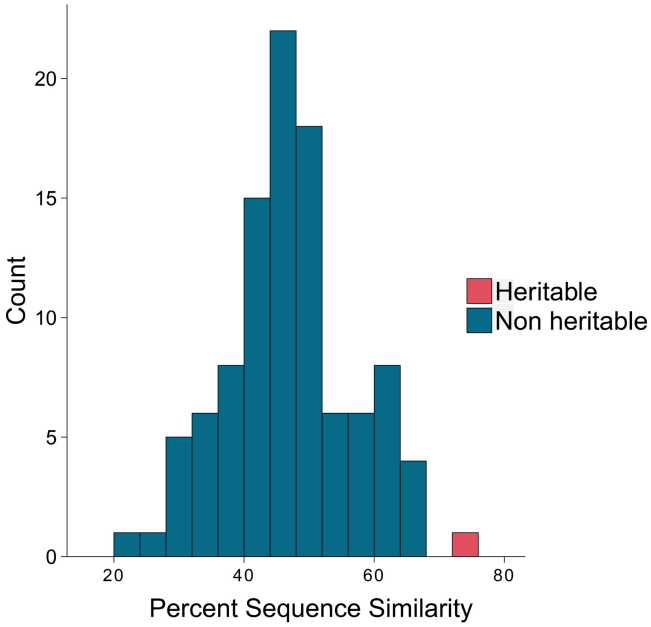

FIG 1 The heritable portion of the bovine rumen microbiome is phylogenetically closely related. The mean pairwise similarity in the 16S rRNA gene sequence of randomly selected groups of rumen OTUs of the same size ($n = 22$) was compared to the mean pairwise similarity of the 22 heritable OTUs. The *y* axis represents the number of groups, and the *x* axis represents the sequence similarity. The group of heritable OTUs with a calculated mean similarity of 72% at the 16S rRNA gene sequence is depicted in pink. The distribution of randomly selected groups of rumen OTUs is depicted in blue. All random groups showed lower mean 16S rRNA gene similarity ($P < 0.01$).

When we measured the phylogenetic distance between these OTUs, we found that they were highly phylogenetically related on the basis of the similarity of their 16S nucleotide sequences (Fig. 1).

These OTUs belong to the two main phyla of the rumen microbiome, namely, *Bacteroidetes* and *Firmicutes*, and grouped under the two dominant orders in the rumen, *Bacteroidales* and *Clostridiales* (Fig. 2).

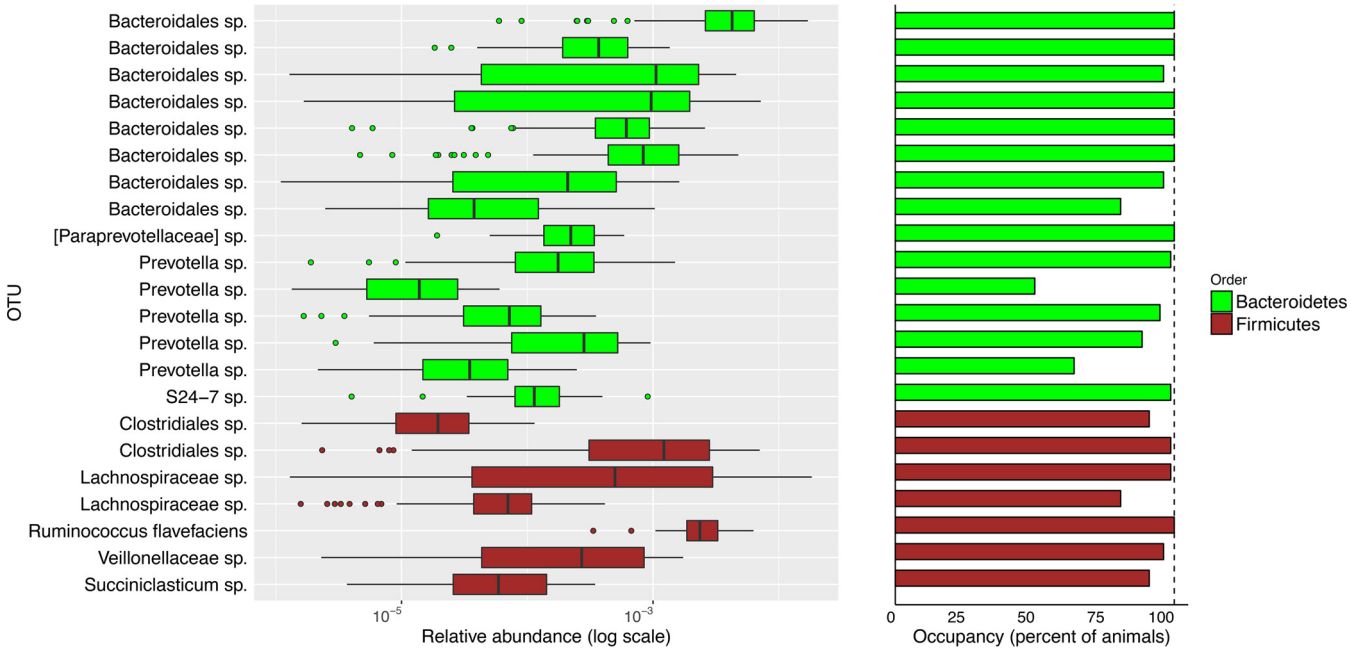

FIG 2 Heritable OTUs show a high presence. OTUs with their taxonomy annotations are listed on the left. The relative abundance of each OTU along the cohort of cows is presented in the left panel, and the presence of each OTU is displayed in the right panel. Green indicates an OTU from the *Bacteroidales* order, while brown indicates an OTU from the *Clostridiales* order.

Downloaded from http://mbio.asm.org/ on June 2, 2019 by guest

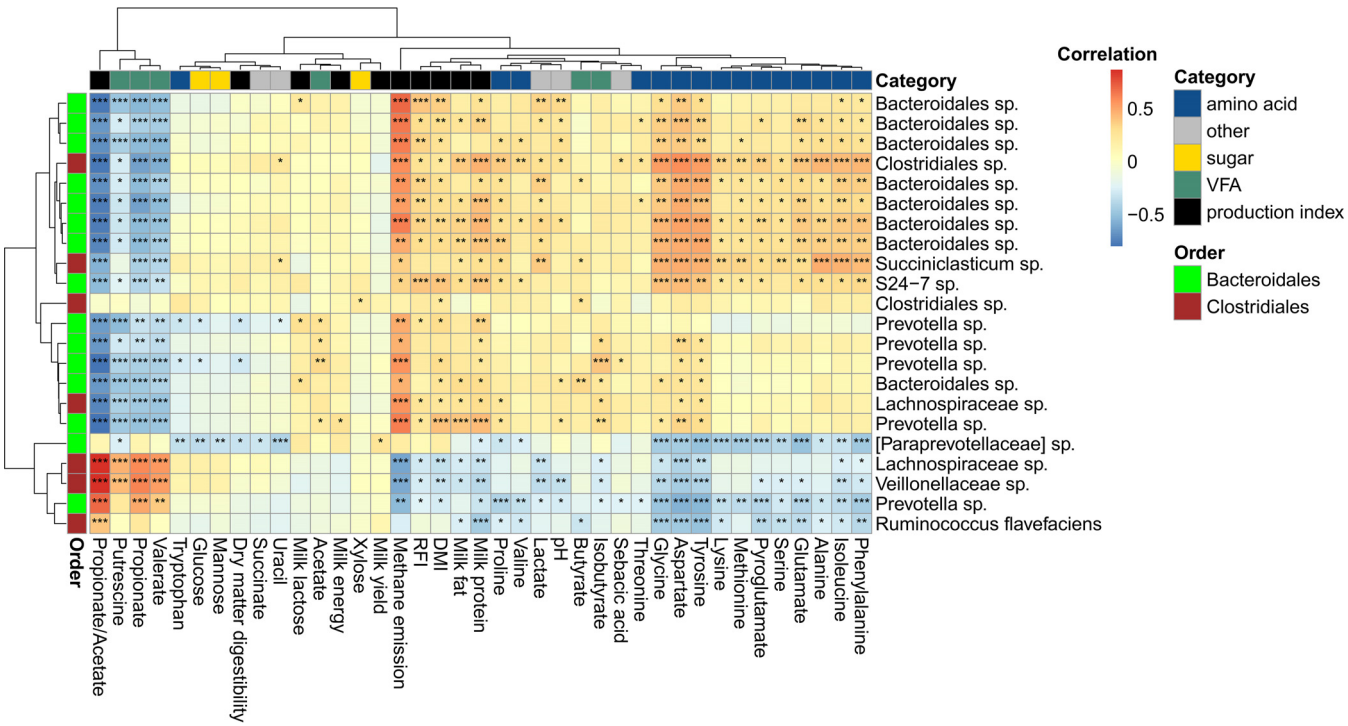

FIG 3 Heritable OTUs are correlated with host attributes and rumen metabolites. A heat map describing the Spearman correlation between the relative abundance of rumen heritable OTUs (rows) and selected indices representing different physiological attributes of the host or rumen metabolites (columns). OTUs are color coded by order (green represents *Bacteroidales*, and brown represents *Clostridiales*). Physiological attributes are colored in black, and rumen metabolites are color coded according to four groups: amino acids (blue), sugars (yellow), VFAs (green), and all other measured metabolites (gray). *, **, and *** represent nominal *P* values smaller than 0.05, 0.005, and 0.0005, respectively.

We further asked whether this phylogenetic composition of heritable OTUs represents that of the overall species composition in the rumen. Here, we found that the order *Bacteroidales* is represented by more species within the heritable OTUs than in the overall rumen microbiome (trend, Fisher exact test, $P < 0.053$).

Heritable bacterial abundance is correlated with host traits as well as with rumen metabolic parameters and can significantly explain a high proportion of the phenotypic variation between animals. Following our recent work in which we found microbiome components that are connected to energy-harvesting efficiency and other physiological parameters of the host (13), we hypothesized that heritable taxa that are correlated with the host genome will potentially be related to rumen metabolism as well as to host physiology. Hence, we looked for a correlation between heritable microbes and all measured physiological parameters of the animals, as well as with rumen metabolic parameters. In detail, we correlated the abundance profile along the cohort of 78 cows of each heritable OTU with the profile of each measured index (a rumen metabolite or other index). We then compared the mean correlation of heritable OTUs with each of the rumen metabolites and host physiological attributes to a null model. In each of 1,000 iterations of the null model, we shuffled each heritable OTU's abundance profile and recalculated its mean correlation with each of the rumen metabolites and host physiological attributes. This analysis revealed that the heritable OTUs exhibit a strong and significant correlation with many of the rumen metabolic parameters, as well with physiological attributes of the host (Fig. 3; Fig. S2).

With relation to rumen metabolism, the strongest correlations for the heritable OTUs were with propionate/acetate ratio (highest-magnitude $r = 0.86$, mean $|r| = 0.64$), methane metabolism (highest-magnitude $r = 0.69$, mean $|r| = 0.49$), propionic acid (highest-magnitude $r = -0.6245274$, mean $|r| = 0.44$), and valeric acid (highest-magnitude $r = -0.57$, mean $|r| = 0.39$), as well as with the concentration of several amino acids, namely, glycine, aspartate, and tyrosine (with highest-magnitude $r = 0.51$, 0.5,

Downloaded from http://mbio.asm.org/ on June 2, 2019 by guest

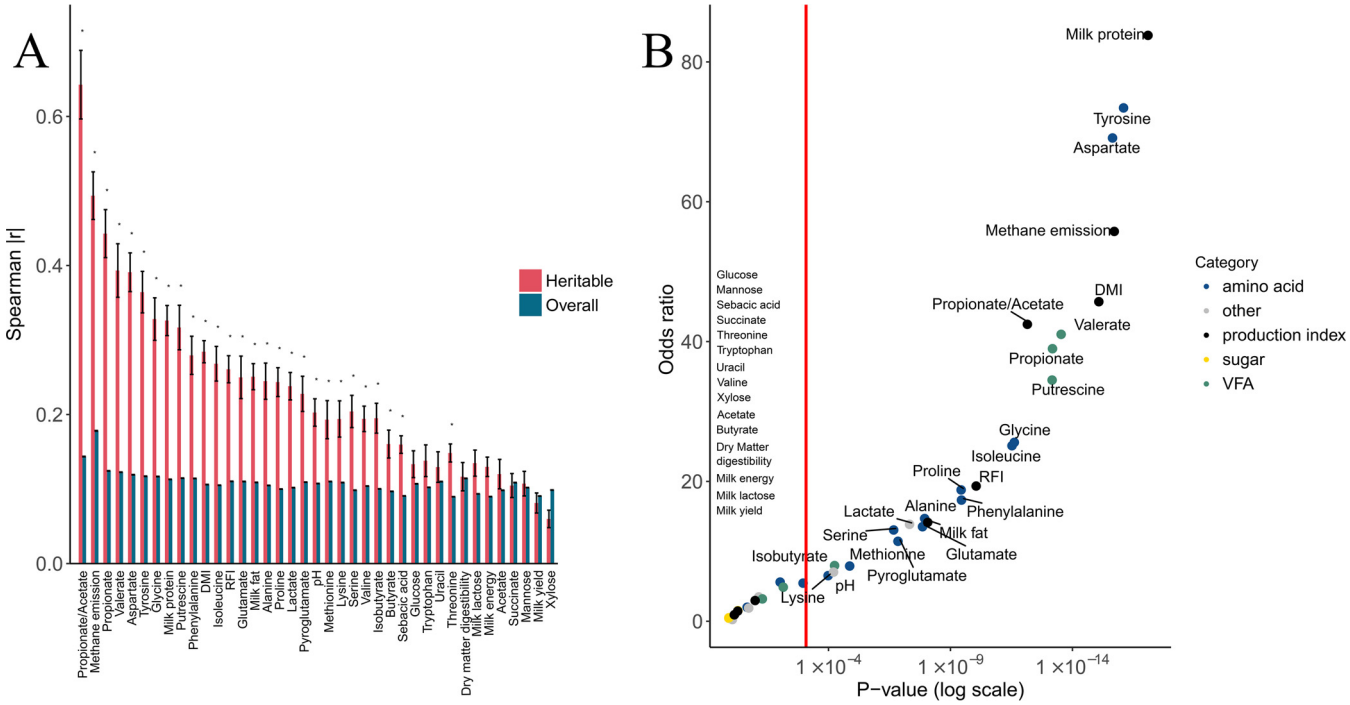

FIG 4 Heritable OTUs are more closely connected to host physiology and rumen metabolites than other rumen microbes. (A) The mean absolute correlation (Spearman) of the heritable OTUs with a given index is compared with that of the entire microbiome. Asterisks represent significant differences in means (*t* test, $P < 0.05$). Red bars represent correlations of the heritable microbiome, while the blue bars represent correlations of the entire microbiome. (B) The odds ratio for an OTU to be correlated with a given index (nominal Spearman *P* of <0.05), between the heritable OTUs and all OTUs. *y* axis, odds ratio; *x* axis, *P* value derived from Fisher's exact test. The red vertical line defines the Bonferroni-corrected 0.05 significance threshold. Point colors signify category according to the legend.

and −0.53 and mean $|r| = 0.32$, 0.39, and 0.36, respectively). Concerning host attributes, the best-correlated parameters were the milk protein (highest-magnitude $r = 0.46$, mean $|r| = 0.33$), dry matter intake (DMI) (highest-magnitude $r = 0.41$, mean $|r| = 0.28$), feed efficiency (represented by residual feed intake [RFI], highest-magnitude $r = 0.26$, mean $|r| = 0.39$), and milk fat (highest-magnitude $r = 0.39$, mean $|r| = 0.25$). Moreover, when we inspected the individual correlation of the heritable OTUs with propionate/acetate ratio, methane metabolism, propionic acid, and valeric acid, the majority of these OTUs were correlated either positively or negatively with these parameters (Fig. 4A). Regarding host physiological attributes, the majority of heritable OTUs were positively correlated with RFI, DMI, and milk protein.

These findings raised the question of whether the portion of heritable microbes that are correlated with host physiology and rumen metabolism is different from the one found in the overall rumen microbiome, as in our recent study we found that the rumen microbiome is tightly linked to many of the host attributes and rumen metabolism parameters (13). To this end, we calculated the OTU correlation odds ratio for each index (see Materials and Methods); in this analysis, we identified significantly higher odds for an OTU to be correlated with a given index within the heritable microbiome for many parameters. This was especially true for the parameters with which these heritable microbes showed high correlation (Fig. 4B).

One of the heritable OTUs with the phylogenetic association of *Bacteroidales*, which was found to be heritable and highly correlated with the feed efficiency trait in this study, was independently found to be correlated with this trait in our previous study (13) (Fig. 5). Additionally, five other heritable OTUs with the phylogenetic associations of *Bacteroidales*, *Prevotella*, *Clostridiales*, and *Flavefaciens* were found to be highly correlated with milk protein in our previous study, and one OTU of the genus *Prevotella* mentioned above was also found to be significantly correlated with dry matter intake (DMI) in our previous study (Fig. 5).

Downloaded from http://mbio.asm.org/ on June 2, 2019 by guest

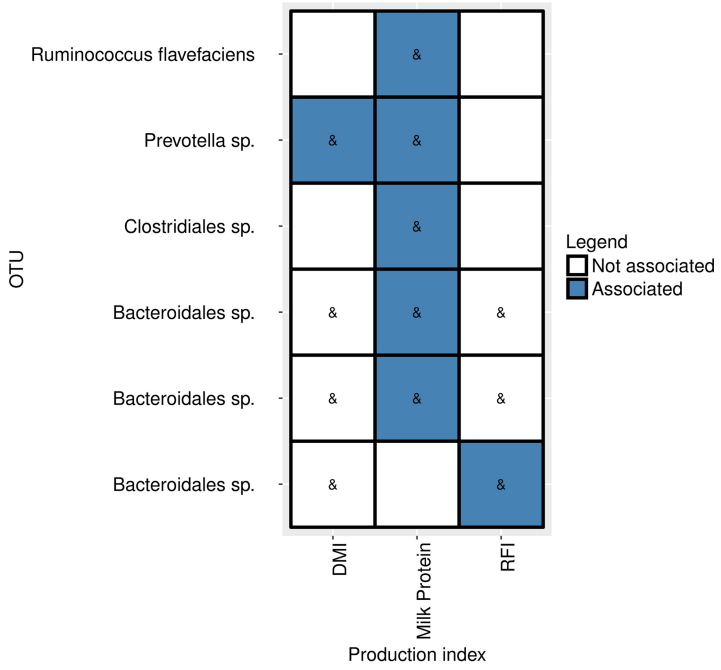

FIG 5 A portion of heritable OTUs were found to be associated with host physiology in a previous study (13). Six of the 22 heritable OTUs that were associated in our previous study with different cow production indices, namely, dry matter intake (DMI), milk protein, and feed efficiency, measured as residual feed intake (RFI). OTUs and their taxonomy are in the rows, and the production indices are in the columns. The ampersand inside a tile indicates that a significant correlation was found, in the current study, between the heritable OTU and the production index.

Rumen and animal physiology traits show various heritability estimates. After identifying heritable microbial species that exhibit correlation with host traits, we were interested in estimating the heritability of the different important host and rumen metabolism traits with which we found the heritable microbes to be correlated (Fig. S6). The volatile fatty acids (VFAs) propionate, succinate, and valerate along with milk protein with the efficiency measures of RFI and DMI exhibited significant heritability estimates. None of the other traits exhibited significant heritability, suggesting that if the cohort size were increased, this trait might prove to be significantly heritable. Alternatively, the current small sample size in our study could also explain the higher estimates for RFI (25) and milk protein (26) than previously published estimates due to intrinsic confounders with the current way of analysis.

DISCUSSION

Our study underlines that the bovine rumen microbiome includes heritable components. We increased resolution over previous studies by applying SNP-based heritability estimates, combined with amplicon sequencing data, host traits, and rumen metabolites. Although we set our threshold for examining microbes that exist in at least 12 genotyped animals (16% of the study cohort), the microbes that were found to be heritable exhibited a considerably higher presence, ranging between 50 and 100% of the animals. This finding may suggest that these microbes are extremely important to rumen metabolism and therefore also to host physiology. These relationships to the host's metabolism and physiology may explain the potential association of these microbial species with the host's genome. However, it should still be noted that in humans and mouse models, microbial taxa showing high narrow sense heritability (h^2) do not necessarily lead to significant associations by quantitative trait locus (QTL) analysis or genome-wide association study (GWAS) (27). Whether this is due to simply missing heritability or to false discovery remains to be determined.

Downloaded from http://mbio.asm.org/ on June 2, 2019 by guest

Our results show that the heritable microbial species represent a related phylogenetic group (Fig. 1). This finding corresponds with a fundamental ecological notion that organisms that share a similar ecological niche are more prone to be phylogenetically close to each other than organisms not sharing the same niche (28). Hence, it is tempting to speculate that the apparent relatedness of the heritable microbial species suggests that they occupy similar ecological niches within the host and potentially share mechanisms of interaction with it, which are affiliated with their phylogenetic association (Fig. 2).

We observed that the metabolites and physiological parameters measured are generally clustered together according to their category, based on hierarchical clustering of their correlation profile for the different heritable OTUs (columns in the heat map [Fig. 3]). For example, most amino acids cluster together, some volatile fatty acids pair together, and six out of nine production indices neighbor each other along the heat map. At the same time, from the heritable OTU perspective, even within the distinct niche of heritable OTUs identified in this study, one can see that the clustering of OTUs according to their abundance profiles (rows in the heat map, Fig. 3) separates them to a high degree according to their taxonomic affiliations, e.g., eight of nine unknown *Bacteroidales* OTUs cluster together and three of five *Prevotella* OTUs cluster together. This finding corresponds with our previous study where we show that specific microbial lineages are correlated with specific physiological traits of the host (8).

Some of the heritable bacteria that were found to be correlated with specific host traits, notably DMI, energy-harvesting efficiency (RFI), and milk protein, were also found independently in our recent study to be connected to these traits, further strengthening these findings (Fig. 5). Intriguingly, we found that the heritable bacteria contain higher proportions of microbes correlated with host traits and with rumen metabolic parameters (Fig. 4A and B). These findings suggest that host genetic variation can have a measurable impact on physiological traits of the host as well as on rumen metabolism by potentially modulating the abundances of different groups of rumen microbes. These findings indicate that host genetics are associated with specific rumen bacteria, which are potentially more prone to influence rumen metabolism and host physiology. Notably, the metabolites and host traits that were found to be correlated with heritable bacteria were also connected by their metabolism. This could be seen in the correlation values of methane production, propionate/acetate ratio, lactate, propionate, and butyrate, as well as energy-harvesting efficiency of the host (represented as RFI), which are correlated with the heritable bacteria. These metabolites were previously shown to be connected to each other by their metabolic pathways (29–31), and the balance between them could affect energy-harvesting efficiency, as was shown in our recent study (13). It is specifically interesting to see that the heritable bacteria are mostly correlated with the propionate/acetate ratio (mean $|r| = 0.64$), which is inversely correlated with methanogenesis and lactate while being positively correlated with RFI, which estimates energy-harvesting efficiency (Fig. 3 and 4). These findings add the host genotype as another dimension to our published findings (13).

In that study, we showed that the metabolism of lactate into VFAs is connected to methane production and increase in energy-harvesting efficiency. These findings were also recently supported by an independent study (32). Another noteworthy finding is the milk protein trait which was correlated with heritable microbes and exhibited the highest odds ratio, pointing to enrichment of heritable bacteria connected to this host trait (Fig. 4). This connection could potentially be explained by the several amino acid rumen concentrations that were also linked to the heritable bacteria, as it was shown elsewhere that increases in amino acids that are not anabolized in the rumen are connected to increases in milk protein (33). Our observations further strengthen the notion of a triangular relationship among the host genotype, rumen bacteria, and host traits. Although it is tempting to speculate that host genetics mediate control of physiological attributes via rumen metabolism, the relationship between these param-

Downloaded from http://mbio.asm.org/ on June 2, 2019 by guest

eters is still to be determined. To target these concerns, future experiments with a larger sample size should be performed to obtain more accurate heritability estimations for both microbes and host traits. Such accuracies could then be utilized to compare heritability estimates of microbes and host traits and suggest cases of the direct causal roles in cases of equal scores. A more direct approach to tackle the question of causality would be to apply transplantation of rumen microbes in different host genetic backgrounds.

Our study presented here sheds light on yet-unanswered questions regarding host microbiome interactions and highlights that host genetic variation is associated with specific microbes. It presents an intricate, unsolved relationship between host genetics, specific microbes, rumen metabolism, and host attributes that will be deciphered by future research involving larger animal cohorts and experimental setups that will allow distinguishing of cause and effect.

MATERIALS AND METHODS

Microbial DNA extraction. The microbial fraction of the rumen fluid was separated according to the method of Stevenson and Weimer (34), with the minor modifications that were performed by Jami et al. (35). DNA was extracted as described by Stevenson and Weimer (34).

Genomic DNA extraction. Five hundred microliters of whole blood from each individual animal was mixed with 500 μl Tris-HCl-saturated phenol (pH 8.0) and 500 μl of double-distilled water (DDW). The mixture was shaken for 4 h at room temperature and subsequently centrifuged at 7,500 $\times$ *g* for 5 min, and the aqueous phase was transferred to a new tube with 500 μl of Tris-HCl (pH 8.0)-saturated phenol-chloroform (1:1) and subsequently centrifuged at 7,500 $\times$ *g* for 5 min. The aqueous phase containing the DNA was transferred to a new tube for further processing.

Animal genotyping. The animals are members of the Volcani Center herd of the Agricultural Research Organization, Israel. Within the genotyped animals, there were 11 groups each sharing a common sire (groups of half-siblings). One such group consisted of four half-siblings, another consisted of three half-siblings, and all the rest of the the groups consisted of two half-siblings. Additionally, there were two pairs of half-siblings sharing a common dam. There were no full siblings among the genotyped animals. As a DNA microsatellite-assisted survey of incorrect paternity attribution within the Israeli dairy cattle population revealed that such incorrect attributions are not rare (36), we opted to base our estimation of genetic relatedness between the cows solely on their genomic information.

Genomic DNA extracts from the animals were loaded into a bovine SNP 50K chip, which is targeted at 54,609 common SNPs that are evenly spaced along the bovine genome (Illumina). The SNP chip model used was Illumina bovine SNP50-24 v3.0, catalog no. 20000766, and it was processed according to the manufacturer's protocol (37) at the Genomics Center of the Biomedical Core Facility, Technion, Israel.

16S rRNA gene sequencing and analysis. Amplification of the 16S V2 region was performed with primers CCTACGGGAGGCAGCAG (forward) and CCGTCAATTCMTTTRAGT (reverse). The libraries were then pooled and subsequently sequenced on a single MiSeq flow cell (Illumina) for 251 cycles from each end of the fragments, following analysis with Casava 1.8. A total of 49,760,478 paired-end reads were obtained from the total sample, with an average of 106,325 paired-end reads per sample. The QIIME (23) pipeline version 1.7.037 was used for data quality control and analyses. OTU analysis was performed on species clusters (97% identity) that were created using UCLUST (38). OTUs were subjected to taxonomy assignment using BLAST (39) against the 16S rRNA reference database RDP (40). Singletons and doubletons were filtered from the data set, resulting in 85,255 species with an average of 5,039 per sample.

Genotype data quality control. Genotypes of 47 individuals from the current analysis were combined with a reference set of 2,691 individual genotypes that were collected from individual Holstein-Friesian dairy cows in farms all over Israel and the Netherlands (courtesy of the Israeli Cow Breeders Association [ICBA]). The reference set of genotypes allowed for more robust quality control (QC) and for the creation of the generic relationship matrix. QC was performed with the PLINK (41) program, with the following parameters: *-cow—file isgenotype_all—maf 0.05— geno 0.05—mind 0.05 –out isgenotype_all_qc—recode12*. SNPs that were not genotyped in more than 5% of the individuals were removed. Similarly, individuals were removed from the analysis if they had been genotyped in less than 95% of the loci (SNPs) covered by the SNP chip.

Three hundred fifty-four individuals (1 belonging to the study group) were removed because of low genotyping, 3,797 SNPs were removed because of "missingness" in the genotyped populations, and 11,290 SNPs failed the minor allele frequency (MAF) criteria. The total number of SNPs passing QC was 40,812.

Generation of genetic relatedness matrix. All animals and SNPs that passed QC were used to generate a matrix that estimates the genetic relatedness between each unique pair of animals. The GCTA (42) software was used to calculate the relationship matrix. The matrix is based on the count of shared alleles, weighted by the allele's rareness:

$$A_{jk} = \frac{1}{n}\sum_{i=1}^{n}\left(\frac{(x_{ij} - 2p_i)(x_{ik} - 2p_i)}{2p_i(1 - p_i)}\right)$$

Downloaded from http://mbio.asm.org/ on June 2, 2019 by guest

where A_{jk} represents the genetic relationship estimate between animals j and k; x_{ij} and x_{ik} are the counts of the reference alleles in animals j and k, respectively; p_i is the proportion of the reference allele in the population; and n is the total number of SNPs used for the relatedness estimation.

Heritability estimates. Heritability estimates of each species were established upon the distribution of the relative abundance of the species in question in conjunction with the estimated genetic relatedness between the animals. The estimation was performed using the software GCTA (42, 43). The model used by this software is termed total heritability and reflects the heritability explained by all the SNPs that passed QC. The model is $y = X\beta + W\mathrm{u} + \varepsilon$, where y is the vector of observations (phenotypes), β is a vector of fixed effects (study covariates), X is the design matrix, u is the vector of SNP effect, W is the standardized genotype matrix, and ε is the individual (residual) effect.

Then, the variance in the model could be attributed to two sources, genetic and random error, in the following manner:

$$V = WW'\sigma_{\mathrm{u}}^2 + I\sigma_{\varepsilon}^2$$

where V is the overall variance, I is the identity matrix ($n \times n$), σ_{u}^2 is the variance due to genetics (overall SNP effects), and σ_{ε}^2 is the variance due to individual effects (residual). Next, GCTA estimates the value of σ_{u}^2 and σ_{g}^2, and the heritability is then estimated as

$$h^2 = \frac{\sigma_{\mathrm{u}}^2}{\sigma_{\mathrm{u}}^2 + \sigma_{\varepsilon}^2}$$

Comparing phylogenetic distance within heritable bacterial OTUs to that within overall rumen microbiome. DNA similarity (percent) between each unique pair within the 22 heritable bacterial OTUs was calculated using Clustal W v2 (44), and the mean of these similarities was then calculated. A reference range of mean similarities was calculated by randomly sampling 100 subsets of the same size, each ($n = 22$) drawn from the pool of OTUs appearing in at least 12 genotyped animals (9,282). Pairwise DNA similarities and their means were calculated for each random subset. To draw significance for the mean similarity within the group of heritable bacterial OTUs, we ranked its mean similarity within all 100 mean similarity values that were obtained from the random subsets.

OTU correlation odds ratio. The OTU correlation odds ratio is (hc/hn)/(nc/nn), where hc is the count of heritable OTUs correlated with the index, hn is the count of heritable OTUs not correlated with the index, nc is the count of nonheritable OTUs correlated with the index, and nn is the count of nonheritable OTUs not correlated with the index. In this context, OTU was correlated with the index if it had a nominal Spearman P value of <0.05.

Statistics and plots. Statistical analysis was performed using R (45) software, and plots were produced using the ggplot2 (46) and pheatmap (47) packages.

SUPPLEMENTAL MATERIAL

Supplemental material for this article may be found at https://doi.org/10.1128/mBio.00703-17.

FIG S1, PDF file, 0.1 MB.
FIG S2, PDF file, 0.1 MB.
FIG S3, PDF file, 0.1 MB.
FIG S4, PDF file, 0.1 MB.
FIG S5, PDF file, 0.1 MB.
FIG S6, PDF file, 0.1 MB.
TABLE S1, TXT file, 0.1 MB.
TABLE S2, TXT file, 0.1 MB.

ACKNOWLEDGMENTS

The research described here was supported by grants from the BARD project no. IS-4801-15C, Israel Science Foundation (1313/13); by the European Research Council under the European Union's Horizon 2020 research and innovation program (grant agreement 640384); and by ICA grant 713 02-15-08a.

We thank Joe I. Weller for useful talks.

REFERENCES

1. Reyniers JA. 1959. The pure-culture concept and gnotobiotics. Ann N.Y Acad Sci 78:3–16. https://doi.org/10.1111/j.1749-6632.1959.tb53091.x.
2. Gilbert JA, Neufeld JD. 2014. Life in a world without microbes. PLoS Biol 12:e1002020. https://doi.org/10.1371/journal.pbio.1002020.
3. Mizrahi I. 2013. Rumen symbioses, p 533–544. *In* Rosenberg E, DeLong EF, Lory S, Stackebrandt E, Thompson F (ed), The prokaryotes: prokaryotic biology and symbiotic associations. Springer, Berlin, Germany.
4. Kittelmann S, Pinares-Patiño CS, Seedorf H, Kirk MR, Ganesh S, McEwan JC, Janssen PH. 2014. Two different bacterial community types are linked with the low-methane emission trait in sheep. PLoS One 9:e103171. https://doi.org/10.1371/journal.pone.0103171.
5. Shi W, Moon CD, Leahy SC, Kang D, Froula J, Kittelmann S, Fan C, Deutsch S, Gagic D, Seedorf H, Kelly WJ, Atua R, Sang C, Soni P, Li D, Pinares-Patiño CS, McEwan JC, Janssen PH, Chen F, Visel A, Wang Z, Attwood GT, Rubin EM. 2014. Methane yield phenotypes linked to differential gene expression in the sheep rumen microbiome. Genome Res 24:1517–1525. https://doi.org/10.1101/gr.168245.113.
6. Ross EM, Moate PJ, Marett LC, Cocks BG, Hayes BJ. 2013. Metagenomic

Downloaded from http://mbio.asm.org/ on June 2, 2019 by guest

predictions: from microbiome to complex health and environmental phenotypes in humans and cattle. PLoS One 8:e73056. https://doi.org/10.1371/journal.pone.0073056.
7. Wallace RJ, Rooke JA, McKain N, Duthie CA, Hyslop JJ, Ross DW, Waterhouse A, Watson M, Roehe R. 2015. The rumen microbial metagenome associated with high methane production in cattle. BMC Genomics 16:839. https://doi.org/10.1186/s12864-015-2032-0.
8. Jami E, White BA, Mizrahi I. 2014. Potential role of the bovine rumen microbiome in modulating milk composition and feed efficiency. PLoS One 9:e85423. https://doi.org/10.1371/journal.pone.0085423.
9. Carberry CA, Kenny DA, Han S, McCabe MS, Waters SM. 2012. Effect of phenotypic residual feed intake and dietary forage content on the rumen microbial community of beef cattle. Appl Environ Microbiol 78:4949–4958. https://doi.org/10.1128/AEM.07759-11.
10. Mizrahi I. 2011. The role of the rumen microbiota in determining the feed efficiency of dairy cows, p 203–210. *In* Rosenberg E, Gophna U (ed), Beneficial microorganisms in multicellular life forms. Springer, Berlin, Germany.
11. Guan LL, Nkrumah JD, Basarab JA, Moore SS. 2008. Linkage of microbial ecology to phenotype: correlation of rumen microbial ecology to cattle's feed efficiency. FEMS Microbiol Lett 288:85–91. https://doi.org/10.1111/j.1574-6968.2008.01343.x.
12. McCann JC, Wiley LM, Forbes TD, Rouquette FM, Jr, Tedeschi LO. 2014. Relationship between the rumen microbiome and residual feed intake-efficiency of Brahman bulls stocked on Bermuda grass pastures. PLoS One 9:e91864. https://doi.org/10.1371/journal.pone.0091864.
13. Shabat SK, Sasson G, Doron-Faigenboim A, Durman T, Yaacoby S, Berg Miller ME, White BA, Shterzer N, Mizrahi I. 2016. Specific microbiome-dependent mechanisms underlie the energy harvest efficiency of ruminants. ISME J 10:2958–2972. https://doi.org/10.1038/ismej.2016.62.
14. Fernando SC, Purvis HT, Najar FZ, Sukharnikov LO, Krehbiel CR, Nagaraja TG, Roe BA, DeSilva U. 2010. Rumen microbial population dynamics during adaptation to a high-grain diet. Appl Environ Microbiol 76:7482–7490. https://doi.org/10.1128/AEM.00388-10.
15. Kim M, Morrison M, Yu Z. 2011. Phylogenetic diversity of bacterial communities in bovine rumen as affected by diets and microenvironments. Folia Microbiol 56:453–458. https://doi.org/10.1007/s12223-011-0066-5.
16. Kittelmann S, Janssen PH. 2011. Characterization of rumen ciliate community composition in domestic sheep, deer, and cattle, feeding on varying diets, by means of PCR-DGGE and clone libraries. FEMS Microbiol Ecol 75:468–481. https://doi.org/10.1111/j.1574-6941.2010.01022.x.
17. Kocherginskaya SA, Aminov RI, White BA. 2001. Analysis of the rumen bacterial diversity under two different diet conditions using denaturing gradient gel electrophoresis, random sequencing, and statistical ecology approaches. Anaerobe 7:119–134. https://doi.org/10.1006/anae.2001.0378.
18. Tajima K, Aminov RI, Nagamine T, Matsui H, Nakamura M, Benno Y. 2001. Diet-dependent shifts in the bacterial population of the rumen revealed with real-time PCR. Appl Environ Microbiol 67:2766–2774. https://doi.org/10.1128/AEM.67.6.2766-2774.2001.
19. Friedman N, Shriker E, Gold B, Durman T, Zarecki R, Ruppin E, Mizrahi I. 2017. Diet-induced changes of redox potential underlie compositional shifts in the rumen archaeal community. Environ Microbiol 19:174–184. https://doi.org/10.1111/1462-2920.13551.
20. Shaani Y, Eliyahu D, Mizrahi I, Yosef E, Ben-Meir Y, Nikbachat M, Solomon R, Mabjeesh SJ, Miron J. 2016. Effect of feeding ensiled mixture of pomegranate pulp and drier feeds on digestibility and milk performance in dairy cows. J Dairy Res 83:35–41. https://doi.org/10.1017/S0022029915000618.
21. Roehe R, Dewhurst RJ, Duthie CA, Rooke JA, McKain N, Ross DW, Hyslop JJ, Waterhouse A, Freeman TC, Watson M, Wallace RJ. 2016. Bovine host genetic variation influences rumen microbial methane production with best selection criterion for low methane emitting and efficiently feed converting hosts based on metagenomic gene abundance. PLoS Genet 12:e1005846. https://doi.org/10.1371/journal.pgen.1005846.
22. Li Z, Wright AG, Si H, Wang X, Qian W, Zhang Z, Li G. 2016. Changes in the rumen microbiome and metabolites reveal the effect of host genetics on hybrid crosses. Environ Microbiol Rep 8:1016–1023. https://doi.org/10.1111/1758-2229.12482.
23. Caporaso JG, Kuczynski J, Stombaugh J, Bittinger K, Bushman FD, Costello EK, Fierer N, Peña AG, Goodrich JK, Gordon JI, Huttley GA, Kelley ST, Knights D, Koenig JE, Ley RE, Lozupone CA, McDonald D, Muegge BD, Pirrung M, Reeder J, Sevinsky JR, Turnbaugh PJ, Walters WA, Widmann J, Yatsunenko T, Zaneveld J, Knight R. 2010. QIIME allows analysis of high-throughput community sequencing data. Nat Methods 7:335–336. https://doi.org/10.1038/nmeth.f.303.
24. Benson AK, Kelly SA, Legge R, Ma F, Low SJ, Kim J, Zhang M, Oh PL, Nehrenberg D, Hua K, Kachman SD, Moriyama EN, Walter J, Peterson DA, Pomp D. 2010. Individuality in gut microbiota composition is a complex polygenic trait shaped by multiple environmental and host genetic factors. Proc Natl Acad Sci U S A 107:18933–18938. https://doi.org/10.1073/pnas.1007028107.
25. VanRaden PM, Van Tassell CP, Wiggans GR, Sonstegard TS, Schnabel RD, Taylor JF, Schenkel FS. 2009. Invited review: reliability of genomic predictions for North American Holstein bulls. J Dairy Sci 92:16–24. https://doi.org/10.3168/jds.2008-1514.
26. Williams YJ, Pryce JE, Grainger C, Wales WJ, Linden N, Porker M, Hayes BJ. 2011. Variation in residual feed intake in Holstein-Friesian dairy heifers in southern Australia. J Dairy Sci 94:4715–4725. https://doi.org/10.3168/jds.2010-4015.
27. Benson AK. 2016. The gut microbiome—an emerging complex trait. Nat Genet 48:1301–1302. https://doi.org/10.1038/ng.3707.
28. Felsenstein J. 1985. Phylogenies and the comparative method. Am Nat 125:1–15.
29. Ungerfeld EM, Rust SR, Burnett R. 2003. Use of some novel alternative electron sinks to inhibit ruminal methanogenesis. Reprod Nutr Dev 43:189–202. https://doi.org/10.1051/rnd:2003016.
30. Öztürk M. 1991. Conversion of acetate, propionate and butyrate to methane under thermophilic conditions in batch reactors. Water Res 25:1509–1513. https://doi.org/10.1016/0043-1354(91)90181-O.
31. Bryant MP. 1979. Microbial methane production—theoretical aspects. J Anim Sci 48:193–201. https://doi.org/10.2527/jas1979.481193x.
32. Kamke J, Kittelmann S, Soni P, Li Y, Tavendale M, Ganesh S, Janssen PH, Shi W, Froula J, Rubin EM, Attwood GT. 2016. Rumen metagenome and metatranscriptome analyses of low methane yield sheep reveals a Sharpea-enriched microbiome characterised by lactic acid formation and utilisation. Microbiome 4:56. https://doi.org/10.1186/s40168-016-0201-2.
33. Rogers JA, Peirce-Sandner SB, Papas AM, Polan CE, Sniffen CJ, Muscato TV, Staples CR, Clark JH. 1989. Production responses of dairy cows fed various amounts of rumen-protected methionine and lysine. J Dairy Sci 72:1800–1817. https://doi.org/10.3168/jds.S0022-0302(89)79297-3.
34. Stevenson DM, Weimer PJ. 2007. Dominance of Prevotella and low abundance of classical ruminal bacterial species in the bovine rumen revealed by relative quantification real-time PCR. Appl Microbiol Biotechnol 75:165–174. https://doi.org/10.1007/s00253-006-0802-y.
35. Jami E, Israel A, Kotser A, Mizrahi I. 2013. Exploring the bovine rumen bacterial community from birth to adulthood. ISME J 7:1069–1079. https://doi.org/10.1038/ismej.2013.2.
36. Ron M, Blanc Y, Band M, Ezra E, Weller JI. 1996. Misidentification rate in the Israeli dairy cattle population and its implications for genetic improvement. J Dairy Sci 79:676–681. https://doi.org/10.3168/jds.S0022-0302(96)76413-5.
37. Adler AJ, Wiley GB, Gaffney PM. 2013. Infinium assay for large-scale SNP genotyping applications. J Vis Exp (81):e50683. https://doi.org/10.3791/50683.
38. Edgar RC. 2010. Search and clustering orders of magnitude faster than BLAST. Bioinformatics 26:2460–2461. https://doi.org/10.1093/bioinformatics/btq461.
39. Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ. 1990. Basic local alignment search tool. J Mol Biol 215:403–410. https://doi.org/10.1016/S0022-2836(05)80360-2.
40. Cole JR, Wang Q, Cardenas E, Fish J, Chai B, Farris RJ, Kulam-Syed-Mohideen AS, McGarrell DM, Marsh T, Garrity GM, Tiedje JM. 2009. The Ribosomal Database Project: improved alignments and new tools for rRNA analysis. Nucleic Acids Res 37:D141–D145. https://doi.org/10.1093/nar/gkn879.
41. Purcell S, Neale B, Todd-Brown K, Thomas L, Ferreira MA, Bender D, Maller J, Sklar P, De Bakker PI, Daly MJ, Sham PC. 2007. PLINK: a tool set for whole-genome association and population-based linkage analyses. Am J Hum Genet 81:559–575. https://doi.org/10.1086/519795.
42. Yang J, Lee SH, Goddard ME, Visscher PM. 2011. GCTA: a tool for genome-wide complex trait analysis. Am J Hum Genet 88:76–82. https://doi.org/10.1016/j.ajhg.2010.11.011.
43. Yang J, Benyamin B, McEvoy BP, Gordon S, Henders AK, Nyholt DR, Madden PA, Heath AC, Martin NG, Montgomery GW, Goddard ME, Visscher PM. 2010. Common SNPs explain a large proportion of the

Downloaded from http://mbio.asm.org/ on June 2, 2019 by guest

heritability for human height. Nat Genet 42:565–569. https://doi.org/10.1038/ng.608.
44. Larkin MA, Blackshields G, Brown NP, Chenna R, McGettigan PA, McWilliam H, Valentin F, Wallace IM, Wilm A, Lopez R, Thompson JD, Gibson TJ, Higgins DG. 2007. Clustal W and Clustal X version 2.0. Bioinformatics 23:2947–2948. https://doi.org/10.1093/bioinformatics/btm404.
45. R Development Core Team. 2016. R: a language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria.
46. Wickham H. 2016. ggplot2: elegant graphics for data analysis. Springer, New York, NY.
47. Kolde R. 2015. pheatmap: pretty heatmaps. R package version 1.0.2.

Downloaded from http://mbio.asm.org/ on June 2, 2019 by guest